National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

University of Alberta

The Evolution of Ontogeny in Water Striders

by

Christian Peter Klingenberg

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Systematics and Evolution

Department of Biological Sciences

Edmonton, Alberta

Spring 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-10602-0

Canada

University of Alberta

Library Release Form

Name of Author: Christian Peter Klingenberg

Title of Thesis: The Evolution of Ontogeny in Water Striders

Degree: Doctor of Philosophy

Year this Degree Granted: 1996

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication an other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Christian Peter Klingenberg
1A 9012-112 Street
Edmonton, Alberta T6G 2C5
Canada

15 April 1996

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *The Evolution of Ontogeny in Water Striders* submitted by Christian Peter Klingenberg in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Systematics and Evolution.

John R. Spence
Dr. J. R. Spence

Dr. B. D. E. Chatterton

Robert T. Hardin
Dr. R. T. Hardin

Dr. B. S. Heming

Brian T. Shea
Dr. B. T. Shea

Cindy Strobeck
Dr. C. Strobeck

Date: 1 April 1996

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9
Canada

Phone: (403) 492-3080 Fax: (403) 492-9234
E-mail: cklingen@gpu.srv.ualberta.ca

9 February 1996

Copyright Permission Department
Plenum Publishing Corporation
233 Spring Street
New York, New York 10013
USA

Copyright Permission Request

I am currently preparing my Ph.D. thesis at the University of Alberta. I plan to include the chapter entitled

Multivariate Allometry by Christian Peter Klingenberg

to appear in the forthcoming book *Advances in Morphometrics*, edited by Leslie F. Marcus, Marco Corti, Anna Loy, Gavin Naylor, and Dennis E. Slice, which will be published by Plenum.

Therefore, I ask you for permission to use the entire chapter as one of eight chapters in my Ph.D. thesis. Credit will be given to the source.

Two copies of this request are enclosed. Please return one and retain one for your records.

Sincerely,

Christian Peter Klingenberg

We hereby grant permission for use of the material as specified above.

Signed: .....

For the copyright holder: .....

PERMISSION GRANTED provided that material has appeared in our work without credit to another source; you obtain the consent of the author(s); you credit the original publication; and reproduction is confined to the purpose for which permission is hereby given.

2/23/96

Georgia Prince Date

Office of Rights/Permissions, Plenum Publishing Corp, 233 Spring Street
New York, NY 10013

FEB 29 1996

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9
Canada

Phone: (403) 492-3080 Fax: (403) 492-9234
E-mail: cklingen@gpu.srv.ualberta.ca

20 February 1996

EVOLUTION, Editorial Office
Department of Zoology
Arizona State University
Tempe, Arizona 85287
USA

Copyright Permission Request

I am currently preparing my Ph.D. thesis at the University of Alberta. I plan to include one article published in *Evolution*, and another currently under review, as chapters in my thesis. Therefore, I ask you for permission to use the following article as a chapter in my Ph.D. thesis:

"Heterochrony and Allometry: Lessons from the Water Strider Genus *Limnoporus*" by Christian Peter Klingenberg and John R. Spence, published in *Evolution* 47:1834–1853, 1993.

Further, I ask you for permission to include the following article, in case the it will be accepted for publication in *Evolution*:

"Individual Variation of Ontogenies: A Longitudinal Study of Growth and Timing" currently under review for possible publication in *Evolution* (MS # 95-212). I am aware that such permission does not imply that the article will be published.

Credit will be given to the source.

Two copies of this request are enclosed. Please return one and retain one for your records.

Sincerely,

Christian Peter Klingenberg

We hereby grant permission for use of the material as specified above.

Signed: Annette C. Keeler FEB 29 1996
Managing Editor, Evolution

For the copyright holder:.....

Interdepartmental Correspondence University of Alberta

1 April 1996

To: **Dr. Christian Klingenberg**
Dept. of Biological Sciences
Biosciences Building

From: Dr. John Spence
Department of Biological Sciences

Dear Chris,

Re: Thesis Chapters

I am pleased to have you use our two joint manuscripts as Chapter 4 (Heterochrony and Allometry in the Water Strider Genus *Limnoporus*) and Chapter 7 (On the Role of Body Size for Life History Evolution) of your Ph.D. thesis. I hereby declare that you had the main creative role in the conceptual content of that work. Our interactions toward producing these two papers were both useful and fun for me. May you have similarly productive interactions in the future!

Sincerely,

John R. Spence
Professor of Biological Sciences

Systematic Biology

David Cannatella
Editor, *Systematic Biology*
Texas Memorial Museum
2400 Trinity
University of Texas
Austin, TX 78705

Phone: (512) 471-9776/5302
Fax: (512) 471-9651
Catfish@mail.utexas.edu

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9
Canada

Dear Chris,

You have permission to use the following paper (currently in press in Systematic Biology) as part of your Ph.D. thesis:

Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components

by Christian Peter Klingenberg, Beat E. Neuenschwander, and Bernard D. Flury

Best wishes,

David Cannatella
Editor

Dr. Beat E. Neuenschwander
Swiss Federal Office of Public Health
Division of Epidemiology

CH-3097 Liebefeld, Hess-Str. 27E
phone: ++ 41 31 323 8815
Fax: ++ 41 31 323 8795
e-mail: 101551.1175@compuserve.com

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9
CANADA

Berne, February 28, 1996

Dear Christian

This letter is to confirm that I have no objections to the inclusion of our joint paper

Ontogeny and individual variation: Analysis of patterned covariance matrices with common principal components
by C. P. Klingenberg, B. E. Neuenschwander, and B. D. Flury
(to appear in Systematic Biology)

in your dissertation.

Best regards

B. E. Neuenschwander

February 12, 1996

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9

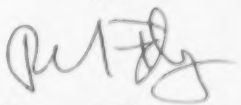
Dear Chris:

This letter is to confirm that I have no objections to the inclusion of our joint paper

Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components, by Klingenberg, Neuenschwander, and Flury, to be published in *Systematic Biology*

in your dissertation. In fact the paper has been almost entirely your work, and I consider it a privilege to be listed as a coauthor.

Sincerely,



Bernard Flury
Professor of Statistics
Department of Mathematics
Indiana University
Rawles Hall
Bloomington, IN 47405
fax: (812) 855-0046
FLURY@INDIANA.EDU

With compliments

Blackwell Science

Blackwell Science Ltd
Osney Mead, Oxford
OX2 0EL, UK
Telephone (01865) 206206
Fax (01865) 721205
http://www.blacksci.co.uk

Christian Peter Klingenberg
Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9
Canada

Phone: (403) 492-3080 Fax: (403) 492-9234
E-mail: cklingen@gpu.srv.ualberta.ca

1 March 1996

Copyright Permission Department
Blackwell Science
Osney Mead
Oxford OX2 0EL
United Kingdom

Copyright Permission Request

I am currently preparing my Ph.D. thesis at the University of Alberta. I plan to include the chapter entitled

On the Role of Body Size for Life History Evolution by Christian Peter Klingenberg and John R. Spence

to appear in *Ecological Entomology*. (The editor, Dr. Simon Letaher, informed me a few days ago that the paper will be accepted for publication.)

Therefore, I ask you for permission to use the entire article as a chapter in my Ph.D. thesis. Credit will be given to the source.

Two copies of this request are enclosed. Please return one and retain one for your records.

Sincerely,

Christian Peter Klingenberg

We hereby grant permission for use of the material as specified above.

Signed: 11/3/96

For the copyright holder: .....

BLACKWELL SCIENCE LTD.
OSNEY MEAD
OXFORD
OX2 0EL

PERMISSION GRANTED, SUBJECT TO THE USUAL ACKNOWLEDGEMENTS OF SOURCE.
BLACKWELL SCIENCE LTD.

I dedicate this thesis

to my parents,
Margrit and Niels Klingenberg,
who awakened my interest in nature
and have always provided encouragement and support,
and

to Rudolf Hauser,
who opened up a new world to me.

Abstract

The link between ontogeny and evolution recently has become the focus of increasing attention by biologists from various fields. My thesis explores the conceptual basis of this connection and provides an empirical case study. First, I review the methods of multivariate allometry and the relationship between heterochrony and allometry. Allometry deals only with morphological traits, whereas the dimension of time is essential for studies of heterochrony—I demonstrate that size cannot be taken as a substitute. I compare the different analytical frameworks of heterochrony, and emphasize that the differences between them can lead to contradictory interpretations about the same data. I apply these methods in a study of the water strider genus *Limnoporus* (Insecta: Heteroptera: Gerridae). Comparisons among the six species of this genus reveals ample flexibility for independent evolution of growth in size and of the durations of ontogenetic stages. A longitudinal study of growth, using a newly developed multivariate method, shows that this evolutionary flexibility corresponds to intraspecific variation in *Limnoporus canaliculatus*. The results of these analyses underscore the difference between incremental and cumulative analyses of growth. Life history is the interface between an organism's ontogeny and its environment; with a large field study in the water strider *Gerris buenoi*, I show that size and developmental time are negatively correlated, not positively, and that female size is unrelated to reproductive performance. These results contradict the prevailing paradigm in life history evolution, but are consistent with the flexibility found in intra- and interspecific variation in *Limnoporus*.

Acknowledgments

I thank John Spence for all his friendship and support. His incessant enthusiasm for our favorite critters and for trying to understand their lives, and his trust and encouragement provided the conditions for me to elvolve—in saltational mode—throughout this part of my ontogeny. Numerous trips to the field and many hours of discussions, in the pond or at the watering hole, made this thesis project feel like pure fun, not work.

The empirical part of this study would not have been possible without the help of many others. I thank Rosalind Barrington Leigh, Nora Berg, Sandy Donald, Shawn Francis, Allen Meyer, and Stacey Rasmussen for invaluable assistance in the field and in the laboratory. Mac McIntyre invented a cheap and effective "pond fart deflection device," which eliminated one of the main perils from the lives of little water striders in my field enclosures. Rosalind Barrington Leigh generously allowed me to use the data for Table 7-2 from a project that is mainly hers.

I owe much of my statistical knowledge to the collaboration with Bernard Flury; he and Beat Neuenschwander developed the model of CPCs for dependent random vectors that is central to chapters 5 and 6, and gave advice on its application. I am very grateful to them.

As members of my supervisory committee, Bruce Heming and Robert Hardin generously provided helpful advice for my project. I thank them and Nils Møller Andersen, Göran Arnqvist, George Ball, Michel Baylac, Bettina Behrens, Fred Bookstein, Brian Chatterton, Marco Corti, Bernard Flury, Laurie Godfrey, Andy Keddie, Carlo Largiadèr, Mättu Loertscher, Simon Leather, Les Marcus, Mike McKinney, Jari Niemelä, Rich Palmer, Locke Rowe, Brian Shea, Curt Strobeck, Kari Vepsäläinen, Miriam Zelditch, Manfred Zimmermann, various participants of the 1993 NATO ASI on "Advances in Morphometrics" and of the Systematics and Evolution discussion group at the University of Alberta, and several anonymous reviewers for various journals for stimulating discussions about my research and for constructive comments on earlier versions of chapters in this thesis.

Financial support for my research was provided by a University of Alberta Ph.D. Scholarship, an Izaak Walton Killam Scholarship, an Andrew Stewart Memorial Graduate Prize to the author, and by grants from the Natural Sciences and Engineering Research Council of Canada to John Spence. Travel support came from NATO and a Mary Louise Imrie Graduate Award (Faculty of Graduate Studies and Research, and the Vice-President [Research] of the University of Alberta).

Table of Contents

1. Introduction **1**
- Background 1
- The Study Organisms 2
- Synopsis 4
- References 5

2. Multivariate Allometry **9**
- Introduction 9
- Levels of Allometry 10
- Principal Components and Allometry 14
- Analyses of Multiple Groups 19
- Example: Geographic Variation 24
- Alternative Approaches in Allometry 29
- References 33
- Appendix 40

4. Heterochrony and Allometry: the Analysis of Evolutionary Change in Ontogeny **41**
- Introduction 41
- Concepts of Heterochrony 43
  - Paedomorphosis and Peramorphosis 43
  - The Clock Model 46
  - The Formalism Based on Growth Functions 49
  - Conflicts Between Different Frameworks 56
    - An Example: Human Heterochrony 57
  - Formalisms Based on Developmental Processes 61
    - Case Studies on the Developmental Basis of Heterochrony 64
- Measures of Time 67
  - Intrinsic and extrinsic time 68
  - Size as Time: "Allometric Heterochrony" 70
- Allometry 73
  - Size, Shape, and Allometry: Two Schools of Thought 74
    - The Huxley–Jolicoeur School 74
    - The Gould–Mosimann School 76
  - Evolution of Ontogenetic trajectories 78
  - Static, Ontogenetic, and Evolutionary Allometry 80
- Conclusions 82
- References 83

4. Heterochrony and Allometry in the Water Strider Genus *Limnoporus* **95**
- Introduction 95
- Heterochrony and Allometry 96
- Materials and Methods 98
  - Study Organisms 98
  - Data 98
  - Statistical Analysis 100
- Results 103
  - Directions of Ontogenetic Trajectories 103
  - Ontogenetic Scaling and Lateral Transposition 104
  - Size Increments 106

Heterochrony 107
Discussion 110
Allometry 111
Heterochrony 113
Conclusions 115
References 116
5. Ontogeny and Individual Variation: Analysis of Patterned Covariance Matrices with Common Principal Components 119
Introduction 119
Statistical Models 120
One-Group Principal Components and Common Principal Components for Independent Groups 121
Longitudinal Data and Patterned Covariance Matrices 122
Common Principal Components for Dependent Random Vectors 124
Example: Growth in Water Striders 126
Data 126
Statistical Analysis and Results 127
Discussion 133
References 135
Appendix 137
Principal Component Analysis 137
Common Principal Components for Independent Groups 138
Common Principal Components for Dependent Random Vectors 138
Number of Parameters 138
Estimation of CPCs 139
6. Individual Variation of Ontogenies: A Longitudinal Study of Growth and Timing 140
Introduction 140
Materials and Methods 142
Laboratory Culture and Measurements 142
Statistical Analyses 143
Results 145
Static Variation: Common Principal Components 145
Relations between Static and Ontogenetic Allometry 148
Individual Variation in Growth 148
Instar Durations 154
Discussion 156
Patterns of Covariation among Morphometric Variables 157
Phenotypic Variation and Constraints in Growth Curves 159
Regulation of Growth 161
Variability in Timing of Molts 162
Possible Physiological Mechanisms 163
Conclusions 164
Literature Cited 165
7. On the Role of Body Size for Life History Evolution 170
Introduction 170
Material and Methods 171
Study Organisms 171

Study Site and Rearing Experiments 171
Breeding Design and Measures of Female Reproduction 173
Estimates of Heritabilities and Genetic Correlations 174
Results 174
Development Time and Final Size 174
Female Size and Reproduction 179
Discussion 183
References 189
8. Synthesis and Outlook 194
References 196

List of Tables

Chapter 2, Appendix. Morphometric data for the example in the text. 40

Table 3-1. Comparison of the terminology for heterochrony proposed by Raff and Wray (1989) with the modified formalism of Alberch et al. (1979). 62

Table 4-1. Ontogenetic allometry of both sexes of the six *Limnoporus* species. 102

Table 4-2. Estimates of the first common principal component (CPC1) and bootstrapped standard errors. 104

Table 4-3. Mean instar durations of the six *Limnoporus* species. 109

Table 5-1. Patterns of individual variation in water striders, analyzed by separate PCAs in each instar. 128

Table 5-2. Joint pattern of variation in all six instars of the water strider data. 129

Table 5-3. Percentages of total within-instar variance taken up by each CPC of the water strider data. 131

Table 6-1. Common principal component coefficients and their bootstrap standard errors. 146

Table 6-2. Percentages of total variance for which the CPCs account within each instar, and their bootstrap standard errors. 147

Table 7-1. Correlations between development time and adult size. 176

Table 7-2. Effect of food limitation and competition on correlations between development time and adult size. 177

Table 7-3. Correlations between female size and reproduction. 180

Table 7-4. Influence of food levels on correlations between female size and reproduction. 183

List of Figures

Fig. 2-1. The three levels of allometry. 11
Fig. 2-2. Types of data for evolutionary allometry. 12
Fig. 2-3. Influence of the correlation between characters. 13
Fig. 2-4. Principal component analysis. 15
Fig. 2-5. Three levels of similarity between covariuance structures. 20
Fig. 2-6. Burnaby's procedure in two dimensions. 22
Fig. 2-7. Geographic variation in the water strider *Gerris costae*. 28
Fig. 2-8. Path models in allometry. 30
Fig. 3-1. The relationship between allometric plots and growth curves. 42
Fig. 3-2. The clock model of Gould (1977). 47
Fig. 3-3. The formalism of Alberch et al. (1979). 51
Fig. 3-4. Effects of pairwise combinations of heterochronic processes. 53
Fig. 3-5. Comparison of growth dynamics in two species. 54
Fig. 3-6. Heterochronic changes and their effects on allometric plots of a trait measurement versus overall size. 72
Fig. 3-7. Evolutionary changes to allometric growth trajectories. 79
Fig. 4-1. Hypothesized phylogeny of the genus *Limnoporus* according to Andersen and Spence (1992). 99
Fig. 4-2. Ordination of multivariate allometric patterns by principal components. 103
Fig. 4-3. Morphometric variation due to lateral transposition of ontogenetic trajectories. 105
Fig. 4-4. Morphometric variation due to non-allometric growth. 106
Fig. 4-5. Size increments. 107
Fig. 4-6. Size increments graphed against instar durations. 108
Figure 4-7. Relationships between the multivariate size measure, age, and instars. 110
Fig. 4-8. Principal components and lateral transposition of growth trajectories. 112
Fig. 5-1. Principal component analysis of a single group. 121
Fig. 5-2. The structure of longitudinal data. 123
Fig. 5-3. The model of common principal components for dependent random vectors. 124
Fig. 5-4. Rearrangement of the patterned covariance matrix after CPC transformation. 125
Fig. 5-5. Covariance matrix for stage-specific measurements in females of the water strider *Limnoporus canaliculatus*. 127

Fig. 5-6. Covariance matrix after transformation to CPCs within each instar. 130
Fig. 5-7. Covariance matrix arranged by CPCs. 132
Fig. 6-1. Comparison of static and ontogenetic allometry. 148
Fig. 6-2. Variation and constraint in instar-specific size. 149
Fig. 6-3. Relations between growth increments of individual water striders in different instars. 150
Fig. 6-4. Patterns of variation in size increments. 151
Fig. 6-5. Relations between relative size in an instar and the size increment in the following molt. 153
Fig. 6-6. Relations between the durations of different instars. 155
Fig. 6-7. Patterns of variation in instar durations. 156
Fig. 6-8. Patterns of variation in cumulative development time. 157
Fig. 6-9. Relations between instar durations and size increments in the following molt. 158
Fig. 7-1. Relation between development time and final size. 175
Fig. 7-2. Relation between female size and measures of reproductive output. 178
Fig. 7-3. Relation between female size and reproductive life span. 181
Fig. 7-4. Relation between female size and measures of egg size and quality. 182
Fig. 7-5. Size-triggered molting as an explanation for the negative correlation between development time and final size. 185
Fig. 7-6. Relationships among life history traits in conventional models and an alternative scenario for *Gerris buenoi*. 188

1. Introduction

Background

The logical connection of individual development to the diversity of species has attracted the interest of researchers since antiquity, even before the idea of evolution gained widespread acceptance (see reviews by Gould 1977; Hall 1992). Yet Charles Darwin's theory of evolution was essential to suggest how this link might have originated, and to initiate widespread research activity.

Early on, scientists interested in evolution discovered the "threefold parallelism" of the diversity of extant species, the progression of individual development, and of geological time as revealed by the fossil record. Darwin's theory immediately suggested a link between diversity and geological time, although the precise nature and scale of "deep time" became known only gradually (Gould 1987). The most comprehensive theory that made use of this connection was Ernst Haeckel's theory of recapitulation (see Gould 1977). The study of ontogeny, Haeckel's term for individual development, was purely descriptive at first, and it was mainly a tool for reconstructing phylogenetic history, based on the *assumption* of recapitulation. The connection between development and evolution remained unclear because of the lack of information regarding both development and its genetic basis.

The study of ontogenetic mechanisms originated as a research program completely separate from any evolutionary considerations. This separation of the new field of experimental and process-oriented developmental biology (Roux's *Entwicklungsmechanik*, i.e., "developmental mechanics") was emphasized in deliberate opposition to the theory of recapitulation, then the generally accepted paradigm, which was becoming untenable in the light of results from embryology and phylogenetic information derived from comparative morphology (of adult forms). For almost a century, the fields of evolutionary and developmental biology remained separate, and only few researchers crossed the disciplinary borderlines.

Evolutionary biologists have dealt with growth mainly through the study of allometry (Huxley 1932; Zuckerman et al. 1950; Cock 1966; Gould 1966). This work generally described patterns, but was not much concerned with the processes from which they originated. During this time, the development of multivariate statistics resulted in new methods for the measurement of the size and shape of organisms (Jolicoeur and Mosimann 1960; Jolicoeur 1963; Burnaby 1966; Hopkins 1966; Mosimann 1970; Sprent 1972). Nevertheless, these developments had little impact on the field, because the new methods were applied only by a relatively small number of biologists.

In the past two decades, however, the connection between development and evolution has been the focus of rapidly increasing interest from both developmental and evolutionary biologists. The most influential contribution was Stephen Jay Gould's book *Ontogeny and Phylogeny* (1977), which reviewed the history of this field and brought the idea of heterochrony to the attention of a broad readership. Gould developed his framework of heterochrony mainly from the work of de Beer (1958). The concept has since been refined and modified in a number of contributions (Alberch et al. 1979; Shea 1983); these

recent developments are reviewed in a series of books (McKinney 1988; McKinney and McNamara 1991; McNamara 1995).

Developmental biologists have also shown a renewed interest in evolution (e.g., Hall 1992; Raff 1992; Hanken 1993; Akam et al. 1994). Molecular methods, which increasingly are being employed to study species other than the classical "model organisms," have produced a large amount of comparative data, and the information from DNA sequences provides a link to molecular systematics. Molecular analyses of development are dealing more and more with evolution, as the history of particular groups of genes is recognized as essential for understanding developmental mechanisms (various contributions in Akam et al. 1994).

Developmental biologists have begun again to use the concept of heterochrony, and they have adapted its analytic frameworks to this new context (Raff and Wray 1989). Interestingly, developmental biologists tend to apply the term heterochrony strictly as a shift in the relative timing of developmental events in the same organism (e.g., Parks et al. 1988, Wray and McClay 1989). This amounts to a resurrection of Haeckel's original meaning of the term, which Gould (1992:165) thought to be "extinct." The renewed use by these authors of another Haeckelian term, heterotopy, is an interesting parallel (Wray and McClay 1989).

Advances in quantitative genetics (e.g., Lande 1979) stimulated new studies on the genetic variation of ontogenetic traits of laboratory rats and mice, and the information on developmental processes made it possible to develop models of development (e.g., Atchley and Rutledge 1980; Cheverud et al. 1983a,b; Atchley 1984, 1987; Leamy and Cheverud 1984; Riska et al. 1984; Slatkin 1987; Atchley and Hall 1991). These endeavors were greatly helped by the rise of a discipline of morphometrics from earlier multivariate studies of organismic form, which has become an increasingly active field of research, as evidenced by a growing number of monographs and symposium proceedings (e.g., Pimentel 1979; Reyment et al. 1984; Bookstein et al. 1985; Rohlf and Bookstein 1990; Bookstein 1991; Reyment 1991; Marcus et al. 1993, 1996).

Altogether, the past two decades have witnessed a gradual approaching of these different disciplines, and a genuine synthesis of developmental and evolutionary biology is emerging (Atkinson 1992; Hall 1992), while at the same time, a similar process of conceptual integration is underway among the various disciplines of morphology (Liem 1991).

This doctoral thesis aims to contribute to this integration of ontogeny and evolution. I mostly focus on the macroscopic aspects of growth and form, and I discuss and use multivariate morphometric methods in an ontogenetic context.

THE STUDY ORGANISMS

Hemimetabolous insects are ideal for the study of ontogeny. They have a richly structured exoskeleton, many parts of which are firmly sclerotized (legs, antennae, etc.). Therefore, it provides excellent opportunities for morphometric measurements. Hemimetabolous insects such as true bugs (Heteroptera) lack the dramatic kind of metamorphosis of holometabolous insects; it is therefore possible to identify morphological landmarks on all structures unambiguously throughout growth.

Because these insects grow by molting, their ontogeny is divided into several discrete steps, which can be easily distinguished. Molts are almost instantaneous compared to the intermolt periods, and therefore allow precise measurement of developmental timing. For instance, in water striders the duration of molts is in the order of 10 minutes, but the instar duration is at least two days and often much longer (pers. obs.).

In groups of species that have a constant number of instars, these constitute homologous developmental stages that can be compared across species. Therefore, the patterns of static, ontogenetic, and evolutionary variation in morphometric traits all can be identified and quantified unambiguously (Klingenberg and Zimmermann 1992a). Instar durations and the morphological characteristics, analyzed in relation to a reconstructed phylogeny of the group, provide an excellent opportunity to study heterochrony.

Many hemimetabolous insects, like the water striders used in this project, can be reared individually in the laboratory. Molting makes these insects particularly valuable subjects for study of growth, because exuviae collected at each molt preserve a complete record of growth of each individual. It is therefore easy to assemble longitudinal data sets without handling the growing larvae at all, thus minimizing possible artifacts. Collected exuviae can be measured later, providing the opportunity to make multiple measurements of structures and to manipulate legs or antennae to align them precisely under the microscope; moreover, it is possible to use high-intensity lighting that would lead to lethal desiccation of larvae. For water striders, individual rearings are even possible in field enclosures, under conditions that quite closely match those of the natural populations (see chapter 7).

Water striders of the family Gerridae are one of the most diverse groups in the hemipteran infraorder Gerromorpha, the semiaquatic bugs (Andersen 1982; Spence and Andersen 1994). They inhabit the surface of a variety of water bodies, ranging from mountain streams to the open sea (see Andersen 1982:272 ff.). In the temperate zone of the Northern Hemisphere, the genera *Gerris* Fabricius, *Limnoporus* Stål, and *Aquarius* Schellenberg are widespread and abundant in many types of aquatic habitats. The phylogeny of these three genera, which form a monophyletic group, is well established through a series of studies by Andersen (1990, 1993) and Andersen and Spence (1992). Moreover, fossils of a *Limnoporus* species from the Middle Eocene of Smithers, British Columbia, establish the minimum age of about 50 million years for the divergence among these genera and of the two principal lineages within *Limnoporus* (Andersen et al. 1993).

All gerrids have five larval instars. These can be easily distinguished, even in the field, because there is a substantial size increase from each instar to the next (e.g., Klingenberg and Zimmermann 1992b), and the instars also differ in coloration and structure (Zimmermann 1987).

The growth and other aspects of the life history of water striders have been the subject of many studies (reviewed by Spence and Andersen 1994). Gerrids are predators and scavengers of insects on the water surface. They overwinter as adults, and produce one or more generations of offspring in the summer months (e.g., Spence 1989). The structure of gerrid communities in different habitats is the result of both active habitat selection by gerrids (Spence 1981) and interactions among species (Spence 1983).

Larval mortality is high in the wild (Zimmermann et al. 1982; Spence 1986a; Klingenberg and Spence 1996a). Field experiments have shown that predation is the dominant mortality cause in *Gerris buenoi*, which is the species studied for the field component of this thesis (Spence 1986a; Klingenberg and Spence 1996a). Important predators include fishing spiders (*Dolomedes*; Zimmermann and Spence 1989) and the larvae of damselflies, dragonflies, and diving beetles (Dytiscidae). Moreover, egg parasitoids can substantially reduce reproductive success of gerrids (Spence 1986b; Nummelin et al. 1988; Henriquez and Spence 1993).

SYNOPSIS

This thesis has two main parts. First, I explore the concepts and methods of multivariate allometry and of heterochrony (chapters 2 and 3). Then, I apply and extend these methods to investigate the evolution of ontogenies in water striders (chapters 4–7).

Multivariate allometry is a branch of morphometrics that is concerned with the covariation of several metric traits associated with variation in overall size of organisms. This approach applies Jolicoeur's (1963) multivariate generalization of Huxley's (1932) equation of simple allometry. This generalization, using principal component analysis, links the study of growth and biological variation to the flexible set of exploratory tools offered by multivariate statistics. In chapter 2, I review this methodology and a number of recent extensions (to be published as Klingenberg 1996). I present an analysis of a small data set from water striders to provide an example of these methods. Finally, I discuss the similarities and differences of this approach to other methods in morphometrics.

Allometry and heterochrony have often been studied and discussed together. This reflects their close relationship, but as a consequence, the clear differences between the two concepts have become blurred, which has led to much confusion in the literature. Therefore, in chapter 3, I review the literature of heterochrony and allometry, emphasizing the differences between these two concepts and between the various methods used in analyses.

As the first section of the empirical part of the thesis chapter I present an application of these methods to study heterochrony and allometry in the water strider genus *Limnoporus* Stål (chapter 4, published as Klingenberg and Spence 1993). This is an interspecific comparison on the background of the phylogeny of the genus (Andersen and Spence 1992), and therefore focuses on the macroevolutionary changes of the ontogenies, on a time scale spanning at least 50 million generations (Andersen et al. 1993).

The microevolutionary processes that underlie these interspecific patterns are the focus of the remaining three chapters. Chapters 5 and 6 use the data from a longitudinal study of phenotypic variation in growth within a single population of *Limnoporus canaliculatus* (Say) from Morris County, New Jersey. I reared these bugs individually in the laboratory, under controlled conditions, and later measured lengths of leg segments from the exuviae. The analysis of such multivariate longitudinal data is difficult, and is tractable only through some simplifying assumptions. The assumption that underlies my analysis is motivated by observations that different developmental stages often

have similar patterns of individual variation. This assumption is incorporated formally in the model of common principal components (CPCs) for dependent random vectors by Flury and Neuenschwander (Neuenschwander 1991; Flury and Neuenschwander 1995). This method is introduced and applied to the data from *L. canaliculatus* in chapter 5 (accepted for publication as Klingenberg et al. 1996). The following chapter uses the components of morphometric variation found in this analysis and examines the patterns of covariation among instars, and the connection between morphometric variation and instar durations. This study identifies constraints on the variability of instar-specific size data, as almost all the phenotypic variation is concentrated in a single dimension. In contrast, a similar analysis of growth increments between instars did not reveal any such constraints. Instar durations are highly variable, but there is a clear correlation between the durations of different instars, demonstrating an even stronger degree of constraint than in the morphometric data.

Studies of life history explore the ecological context of the evolution of ontogenies. Chapter 7 (to be published as Klingenberg and Spence 1996b) reports on a large field study testing the assumptions that underlie models of life history evolution (Roff 1992; Stearns 1992). This study, using unusually large sample sizes in multiply replicated experiments, clearly contradicts two key assumptions of the current paradigm in life history theory. First, development time and adult size are negatively, not positively, correlated and therefore indicate variation in overall vigor, rather than the trade-off assumed by theoretical models. Second, of several measures of female reproductive performance, none is correlated with body size, suggesting that the variation in size found within this population has no consequences for reproductive fitness. Within the bounds of intrapopulation variation, size may thus evolve as a neutral trait.

The four chapters of the empirical part of this thesis examine ontogenetic variation at the interspecific and intrapopulation level, and also explore its ecological context. Altogether, they cover many of the theoretical questions discussed in the theoretical chapters 2 and 3. In chapter 8, I provide a brief discussion of the contribution made in this thesis and an outlook on future work necessary for a full understanding of the evolution of ontogeny in water striders.

References

Akam, M., P. Holland, P. Ingham, and G. Wray, eds. 1994. The evolution of developmental mechanisms. Development 1994 Supplement.

Alberch, P., S. J. Gould, G. F. Oster, D. B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5:296–317.

Andersen, N. M. 1982. The semiaquatic bugs (Hemiptera, Gerromorpha): phylogeny, adaptations, biogeography, and classification. Scandinavian Science Press, Klampenborg, Denmark.

Andersen, N. M. 1990. Phylogeny and taxonomy of water striders, genus *Aquarius* Schellenberg (Insecta, Hemiptera, Gerridae), with a new species from Australia. Steenstrupia 16:37–81.

Andersen, N. M. 1993. Classification, phylogeny, and zoogeography of the pond skater genus *Gerris* Fabricius (Hemiptera: Gerridae). Canadian Journal of Zoology 71:2473–2508.

Andersen, N. M., and J. R. Spence. 1992. Classification and phylogeny of the Holarctic water strider genus *Limnoporus* Stål (Hemiptera, Gerridae). Canadian Journal of Zoology 70:753–785.

Andersen, N. M., J. R. Spence, and M. V. H. Wilson. 1993. 50 Million years of structural stasis in water striders (Hemiptera: Gerridae). American Entomologist 39:174–176.

Atchley, W. R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. American Naturalist 123:519–540.

Atchley, W. R. 1987. Developmental quantitative genetics and the evolution of ontogenies. Evolution 41:316–330.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. Biological Reviews 66:101–157.

Atchley, W. R., and J. J. Rutledge. 1980. Genetic components of size and shape. I. Dynamics of components of phenotypic variability and covariability during ontogeny in the laboratory rat. Evolution 34:1161–1173.

Atkinson, J. W. 1992. Conceptual issues in the reunion of development and evolution. Synthese 91:93–110.

Bookstein, F. L. 1991. Morphometric tools for landmark data: geometry and biology. Cambridge University Press, Cambridge.

Bookstein, F. L., B. Chernoff, R. L. Elder, J. M. Humphries, Jr., G. R. Smith, and R. E. Strauss. 1985. Morphometrics in evolutionary biology: the geometry of size and shape change, with examples from fishes. Academy of Natural Sciences of Philadelphia, Special Publication 15.

Burnaby, T. P. 1966. Growth-invariant discriminant functions and generalized distances. Biometrics 22:96–110.

Cheverud, J. M., L. J. Leamy, W. R. Atchley, and J. J. Rutledge. 1983a. Quantitative genetics and the evolution of ontogeny: I. Ontogenetic changes in quantitative genetic variance components in randombred mice. Genetical Research (Cambridge) 42:65–75

Cheverud, J. M., J. J. Rutledge, W. R. Atchley. 1983b. Quantitative genetics of development: genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905.

Cock, A. G. 1966. Genetical aspects of metrical growth and form in animals. Quarterly Review of Biology 41:131–190.

de Beer, G. 1958. Embryos and ancestors. Oxford University Press, Oxford.

Flury, B. D., and B. E. Neuenschwander. 1995. Principal component models for patterned covariance matrices, with applications to canonical correlation analysis of several sets of variables. Pp. 90–112 *in* W. J. Krzanowski, ed. Recent advances in descriptive multivariate analysis. Oxford University Press, Oxford.

Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biological Reviews 41:587–640.

Gould, S. J. 1977. Ontogeny and phylogeny. Harvard University Press, Cambridge, Mass.

Gould, S. J. 1987. Time's arrow, time's cycle: myth and metaphor in the discovery of geological time. Harvard University Press, Cambridge, Mass.

Gould, S. J. 1992. Heterochrony. Pp. 158–165 *in* E. F. Keller and E. A. Lloyd, eds. Keywords in evolutionary biology. Harvard University Press, Cambridge, Mass.

Hall, B. K. 1992. Evolutionary Developmental Biology. Chapman and Hall, London.

Hanken, J. 1993. Model systems versus outgroups: alternative approaches to the study of head development and evolution. American Zoologist 33:448–456.

Henriquez, N. P., and J. R. Spence. 1993. Host location by the gerrid egg parasitoid *Tiphodytes gerriphagus* (Marchal) (Hymenoptera: Scelionidae). Journal of Insect Behavior 6:455–466.

Hopkins, J. W. 1966. Some considerations in multivariate allometry. Biometrics 22:747–760.

Huxley, J. S. 1932. Problems of relative growth. Methuen, London. Reprinted 1993, Johns Hopkins University Press, Baltimore.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. Biometrics 19:497–499.

Jolicoeur, P., and J. E. Mosimann. 1960. Size and shape variation in the painted turtle: a principal component analysis. Growth 24:339–354.

Klingenberg, C. P. 1996. Multivariate allometry. Pp. 23–49 *in* L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.

Klingenberg, C. P., B. E. Neuenschwander, and B. D. Flury. 1996. Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components. Systematic Biology. In press.

Klingenberg, C. P., and J. R. Spence. 1993. Heterochrony and allometry: lessons from the water strider genus *Limnoporus*. Evolution 47:1834–1853.

Klingenberg, C. P., and J. R. Spence. 1996a. Impacts of predation and intracohort cannibalism in the water strider *Gerris buenoi* (Heteroptera: Gerridae). Oikos 75. In press.

Klingenberg, C. P., and J. R. Spence. 1996b. On the role of body size for life history evolution. Ecological Entomology. In press.

Klingenberg, C. P., and M. Zimmermann. 1992a. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of water striders. American Naturalist 140:601–620.

Klingenberg, C. P., and M. Zimmermann. 1992b. Dyar's rule and multivariate allometric growth in nine species of waterstriders (Heteroptera: Gerridae). Journal of Zoology (London) 227:453–464.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. Evolution 33:402–416.

Leamy, L., and J. M. Cheverud. Quantitative genetics and the evolution of ontogeny. II. Genetic and environmental correlations among age-specific characters in randombred house mice. Growth 48:339–353.

Liem, K. F. 1991. Toward a new morphology: pluralism in research and education. American Zoologist 31:759–767.

Marcus, L. F., E. Bello, and A. García-Valdecasas, eds. 1993. Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Marcus, L. F., M. Corti, A. Loy, G. Naylor, D. E. Slice. 1996. Advances in morphometrics. Plenum Press.

McKinney, M. L., ed. 1988. Heterochrony in evolution: a multidisciplinary approach. Plenum Press, New York.

McKinney, M. L., and K. J. McNamara. 1991. Heterochrony: the evolution of ontogeny. Plenum Press, New York.

McNamara, K. J., ed. 1995. Evolutionary change and heterochrony. John Wiley & Sons, Chichester.

Mosimann, J. E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association 65:930–945.

Neuenschwander, B. 1991. Common principal components for dependent random vectors. Thesis, University of Bern, Switzerland.

Nummelin, M., J. R. Spence, and K. Vepsäläinen. 1988. Infection of gerrid eggs (Heteroptera: Gerridae) by the parasitoid *Tiphodytes gerriphagus* Marchal (Hymenoptera: Scelionidae) in Finland. Annales Zoologici Fennici 25:299–302.

Parks, A. L., B. A. Parr, J.-E. Chin, D. S. Leaf, and R. A. Raff. 1988. Molecular analysis of heterochronic changes in the evolution of direct developing sea urchins. Journal of Evolutionary Biology 1:27–44.

Pimentel, R. A. 1979. Morphometrics: the multivariate analysis of biological data. Kendall/Hunt, Dubuque, Iowa.

Raff, R. A. 1992. Evolution of developmental decisions and morphogenesis: the view from two camps. Development 1992 Supplement:15–22.

Raff, R. A., and G. A. Wray. 1989. Heterochrony: developmental mechanisms and evolutionary results. Journal of Evolutionary Biology 2:409–434.

Reyment, R. A. 1991. Multidimensional palaeobiology. Pergamon Press, Oxford.

Reyment, R. A., R. E. Blackith, and N. A. Campbell. 1984. Multivariate morphometrics. Academic Press, London.

Riska, B., W. R. Atchley, and J. J. Rutledge. 1984. A genetic analysis of targeted growth in mice. Genetics 107:79–101.

Roff, D. A. 1992. The evolution of life histories: theory and analysis. Chapman & Hall, New York.

Rohlf, F. J., and F. L. Bookstein, eds. 1990. Proceedings of the Michigan morphometrics workshop. University of Michigan Museum of Zoology, special Publication No. 2.

Shea, B. T. 1983. Allometry and heterochrony in the African apes. American Journal of Physical Anthropology 62:275–289.

Slatkin, M. 1987. Quantitative genetics of heterochrony. Evolution 41:799–811.

Spence, J. R. 1981. Experimental analysis of microhabitat selection in water-striders (Heteroptera: Gerridae). Ecology 62:1505–1514.

Spence, J. R. 1983. Pattern and process in co-existence of water-striders (Heteroptera: Gerridae). Journal of Animal Ecology 52:497–511.

Spence, J. R. 1986a. Relative impacts of mortality factors in field populations of the water strider *Gerris buenoi* Kirkaldy (Heteroptera: Gerridae). Oecologia (Berlin) 70:68–76.

Spence, J. R. 1986b. Interactions between the scelionid egg parasitoid *Tiphodytes gerriphagus* (Hymenoptera) and its gerrid hosts (Heteroptera). Canadian Journal of Zoology 64:2728–2738.

Spence, J. R. 1989. The habitat templet and life history strategies of pond skaters (Heteroptera: Gerridae): reproductive potential, phenology, and wing dimorphism. Canadian journal of Zoology 67:2432–2447.

Spence, J. R., and N. M. Andersen. 1994. Biology of water striders: interactions between systematics and ecology. Annual Review of Entomology 39:101–128.

Sprent, P. 1972. The mathematics of size and shape. Biometrics 28:23–37.

Stearns, S. C. 1992. The evolution of life histories. Oxford University Press, Oxford.

Wray, G. A., and D. R. McClay. 1989. Molecular heterochronies and heterotopies in early echinoid development. Evolution 43:803–813.

Zimmermann, M. 1987. Die Larven der schweizerischen *Gerris*-Arten. Revue Suisse de Zoologie 94:593–624.

Zimmermann, M., R. Hauser, and J. Hüsler. 1982. Schätzung der stadienspezifischen Dichten und Überlebensraten in einer Larvenpopulation von *Gerris lacustris*. Revue Suisse de Zoologie 89:941–955.

Zimmermann, M., and J. R. Spence. 1989. Prey use of the fishing spider *Dolomedes triton* (Pisauridae, Araneae): an important predator of the neuston community. Oecologia (Berlin) 80:187–194.

Zuckerman, S., G. M. Morant, J. Hammond, F. G. Gregory, J. S. Huxley, D. A. Sholl, P. B. Medawar, F. Yates, J. B. S. Haldane, T. S. Westoll, C. H. Waddington. 1950. A discussion on the measurement of growth and form. Proceedings of the Royal Society of London B 137:433–523.

2. Multivariate Allometry*

Introduction

Variation in size of organisms usually is associated with variation in shape, and most metric characters are highly correlated among one another. These associations are the subject of allometry (Huxley, 1932; Cock, 1966; Gould, 1966, 1975). Although allometry is often used to examine the consequences of size for ecological or physiological variables (Günther, 1975; LaBarbera, 1989; Reiss, 1989), this review deals only with measurements of traits used to characterize the morphological form of organisms.

Unlike other approaches in morphometrics, which are built on geometric theory, allometry has a largely empirical basis. Huxley (1932) realized that scatter plots of two trait measurements in growing organisms often closely follow a curved line, and that this relationship usually becomes linear if both measurements are transformed to logarithms. From this, he derived his *formula of simple allometry* $y = b\,x^{\alpha}$ or, in log-transformed notation, $\log y = \log b + \alpha \log x$, where x and y are trait measurements, and b and α are constants. The constant α, the slope in log-log plots of x and y, is often called the allometric coefficient (terminology is not uniform; some authors call b coefficient). The special case when $\alpha = 1$ is called *isometry*, and indicates direct proportionality between x and y. If $\alpha > 1$, there is *positive allometry*, whereas for *negative allometry*, $\alpha < 1$ (Huxley and Teissier, 1936). In humans, for example, the long bones of the limbs show positive allometric growth relative to overall stature, and the height of the head negative allometry.

In most morphometric data sets, measurements are positively correlated, i.e., x and y increase or decrease simultaneously. Even if there is negative allometry, α still is positive; negative allometry only implies that the relative variation in y is smaller than in x, e.g., if y grows by 10% for every 20% growth increment in x. If α is negative, however, there is an absolute reduction in y associated with an increase in x. This case is called *enantiometry* (Huxley and Teissier, 1936). Reduction of the absolute size of organs during growth is a real phenomenon, although it is not found commonly in morphometric studies. The most striking example is the shrinking of larval structures during metamophosis, e.g. the gills and tail of anuran tadpoles, but in a subtler way, it even occurs in cranial growth of primates (Corner and Richtsmeier, 1991).

Huxley's approach is not restricted to pairs of measurements. In many multivariate data sets, log-log plots of all pairwise combinations of morphometric variables show approximately linear relationships. Therefore, Huxley's bivariate allometry can be generalized to multiple dimensions. Moreover, it is not confined to growth data, as straight-line relationships are also found in log-log plots of intra- and interspecific variation within one particular ontogenetic

---

* A version of this chapter has been accepted for publication. C. P. Klingenberg 1996. Pp. 23–49 in L. F. Marcus, M. Corti, A. Loy, G. Naylor, D. E. Slice (editors), *Advances in Morphometrics*. Plenum Press, New York.

stage (most often adults). From this perspective of allometry, some major questions are: How much variation is there? Are the data points clustered around a straight line, and if so, how closely? What is the direction of that line in multidimensional space? Do different groups of organisms share the same allometric relationship?

The purpose of this chapter is to review concepts and techniques used in studies of multivariate allometry. First, I introduce the main levels of variation within and between species, which have been the subject of allometric studies. Then, I present the multivariate generalization of allometry using principal component analysis and some more recent developments, such as the bootstrap and techniques dealing with multiple groups. Finally, I briefly contrast some alternative approaches to allometry.

LEVELS OF ALLOMETRY

Huxley (1932) devised allometry mainly as a tool to study the relative growth of parts in various organisms. Growth, however, is not the only origin of variation in overall size and associated variation in shape, because evolutionary changes and individual variation also can generate allometric relations. These levels or types of allometry have been included in elaborate classification schemes (see Cock, 1966; Gould, 1966, 1975). Because of its simplicity, I prefer the terminology proposed by Cock (1966) who distinguished static, ontogenetic, and evolutionary allometry (Fig. 2-1). This classification has also been used in most empirical comparisons between levels of allometry (Cheverud, 1982; Leamy and Bradley, 1982; Boag, 1984; Gibson et al., 1984; Leamy and Atchley, 1984*a*; Shea, 1985; Klingenberg and Zimmermann, 1992*a*).

Static allometry, which is also referred to as *size allometry*, results from variation among individuals of the same population and age group (intraspecific scaling, Gould, 1975). Static allometry is particularly easy to study in organisms with discrete growth stages, such as insects (Cuzin-Roudy, 1975; Klingenberg and Zimmermann, 1992*a*), or in adults of organisms with determinate growth, such as birds (Boag, 1984; Gibson et al., 1984). These studies, among others, found that the largest proportion of multivariate variation is contained in one dimension, and that the model of simple static allometry therefore is appropriate. This phenomenon has been termed morphometric or phenotypic integration (e.g., Leamy and Atchley, 1984*b*; Zelditch, 1987, 1988; Zelditch and Carmichael, 1989). Although there is an extensive literature describing static allometry and morphometric integration, relatively little is known about their developmental basis (but see Cowley and Atchley, 1990; Atchley and Hall, 1991; Shea, 1992; Paulsen and Nijhout, 1993). In theoretical models, Riska (1986) explored how developmental processes can affect static correlations among the traits they produce (see also Cowley and Atchley, 1992). Patterns of static allometry have sometimes been used to deduce underlying developmental processes (e.g., Zelditch, 1987; Wheeler, 1991). Such inferences, however, should be substantiated by direct observations.

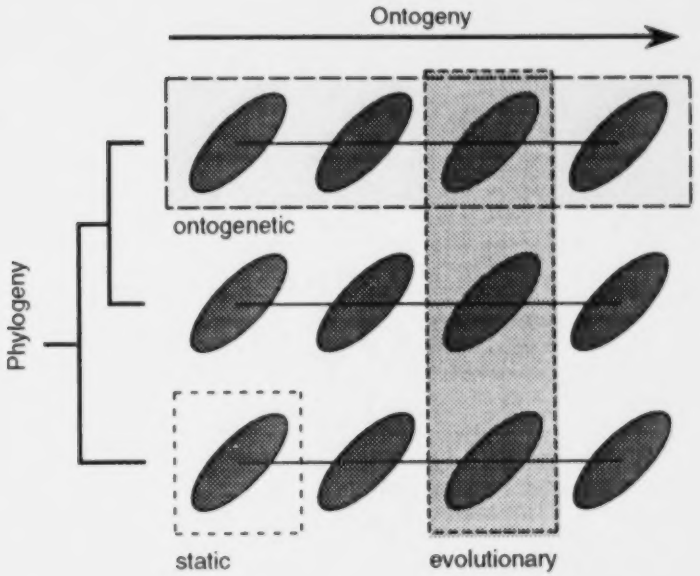


Fig. 2-1. The three levels of allometry. The diagram shows three species, each with four different ontogenetic stages, which are considered to be homologous among species. Rectangles enclose the species and stage groups included in an analysis of allometry at each of the three levels. Ontogenetic allometry can be separately analyzed for all three species, evolutionary allometry for each of the four stages, and static allometry for each of the 12 species and stage groups.

Ontogenetic allometry or *growth allometry* deals with covariation among characters during growth. Simple allometry occurs if the ratio between the specific growth rates (percentage increment per time unit) of two different characters is constant (Huxley, 1932; Reeve and Huxley, 1945; Shea, 1985; Blackstone, 1987). Theoretical studies showed that various models of growth as a function of time can result in simple allometry (e.g., Laird, 1965; Laird et al., 1968; Katz, 1980). The rule of simple allometry holds often, but not always, as Huxley (1932) demonstrated with an impressive list of bivariate examples. Correspondingly, multivariate studies often find that one dimension contains the largest proportion of the total variation, sometimes more than 99% (e.g., Jungers et al., 1988; Solignac et al., 1990; Strauss, 1990*b*; Klingenberg and Zimmermann, 1992*a*). Some studies, however, show clear deviations from simple allometry in certain structures (Cuzin-Roudy and Laval, 1975; Boag, 1984; Cane, 1993), or subtle curvatures of growth trajectories in the space of log-transformed characters (Bookstein, 1991: Figs. 4.2.2 to 4.2.4; Klingenberg and Zimmermann 1992*a*; Klingenberg and Spence, 1993). Studies of plant growth showed particularly strong deviations from simple allometry (e.g. Kampny et al., 1993; McLellan, 1993). Three types of data are used to study ontogenetic allometry: longitudinal data with measurements of the same individuals at several developmental stages, cross-sectional data with different

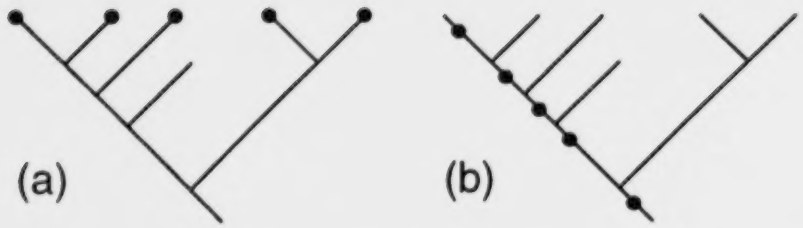


Fig. 2-2. Types of data for evolutionary allometry. (a) In neontological studies, the data always are measurements taken on recent species from several lineages in a clade, which are related as sister groups rather than ancestors and descendants. (b) In paleontological studies, evolutionary allometry often refers to character covariation among members of a single evolutionary lineage. Because it is difficult to distinguish with certainty whether two fossil species are related as ancestor and descendant or as sister groups, many of these studies may in fact use designs that are mixtures of (a) and (b).

specimens in several known stages, and mixed cross-sectional data without information on ontogenetic stage (Cock, 1966).

Evolutionary allometry reflects covariation among changes in different traits along the branches of a phylogeny. It is concerned with character covariation among contemporaneous species sharing a common ancestor (Fig. 2-2a), or among fossil members of an evolutionary lineage (Fig. 2-2b). I do not distinguish separate levels for analyses using these two types of data (for different terminology, see Gould, 1966). Evolutionary processes leading to the associations between trait changes presumably do not differ depending on whether these changes occur within one lineage successively or in different lineages giving rise to sister groups. It is important to use specimens in comparable ontogenetic stages to avoid confounding evolutionary and ontogenetic variation. This is straightforward in organisms with determinate growth, such as birds or insects (e.g., Livezey, 1989; Strauss, 1990*a*; Klingenberg and Zimmermann, 1992*a*), but it is more difficult in organisms with indeterminate growth, where some studies assume that specimens are "typical" for the respective species (e.g., Strauss, 1985). In these studies, among others, the model of simple allometry fits the data fairly well, indicating that evolutionary variation is "constrained" in its dimensionality (Maynard Smith et al., 1985; Gould, 1989; Arnold, 1992). Some of this covariation among traits may be determined by developmental processes, as Riska (1989) showed with a simulation study.

Evolutionary allometry, like all interspecific comparisons, presents some statistical problems because the species are not independent of one another, but are parts of a hierarchically structured phylogeny (e.g., Felsenstein, 1985; Pagel and Harvey, 1988). This interdependence is most evident for species presumed to be members of a single, unbranched lineage (Fig. 2-2b), but it also applies to comparisons among terminal taxa in a clade. A possible solution is the method of phylogenetically independent contrasts (Felsenstein, 1985; Martins and Garland, 1991; Garland et al., 1992), which analyzes character changes along the branches in the phylogeny instead of the measurements (character states) of terminal taxa. Changes are either directly measured, if ac-

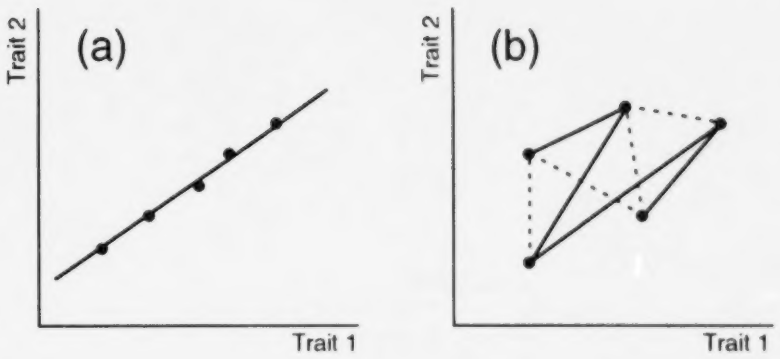


Fig. 2-3. Influence of the correlation between characters on the robustness of evolutionary allometry. For simplicity, I assume that the species are members of a single evolutionary lineage (as in Fig. 2-2b). (a) If the traits are highly correlated, most morphometric variation is in a single direction. Estimates of evolutionary allometry, the "average" direction of evolutionary changes, will yield almost the same result regardless of the phylogenetic relations among species (i.e., how the points are connected). (b) If the correlation between traits is low, however, the "average" direction of changes depends strongly on the phylogenetic scenario, i.e., on how the points are connected. The solid and dashed lines represent two alternative hypotheses of ancestor-descendant relationships, which have drastically different directions of evolutionary changes.

tual ancestor-descendant series of fossils are available, or inferred by indirect methods such as parsimony or maximum likelihood. This approach requires knowledge of the phylogeny of the group under consideration, and a multivariate theory for reconstructing state vectors of quantitative characters for internal nodes, which still needs to be developed (for univariate approaches, see Maddison and Maddison, 1992). For morphometric data, however, the problem of phylogenetic dependence may not be as severe as for other data types, because often most of the variation is in a single dimension. These high correlations make estimates of the direction of evolutionary changes relatively robust against errors in phylogenetic reconstruction (Fig. 2-3).

Further levels of allometry exist in organisms with modular organization, such as colonial animals and most vascular plants. In addition to the ontogeny of individual zooids, colonial animals have an additional level of colony-wide development, which is called astogeny (e.g., Pandolfi, 1988). Buss and Blackstone (1991) showed that colony growth in a marine hydroid follows well-determined trajectories, and that colonies react to experimental perturbations in an integrated manner (see also Anstey, 1987). Similarly, the structure of plant parts changes with the age of the entire plant (heteroblasty; e.g., McLellan, 1993). Jones (1992, 1993) studied correspondences between the development of individual leaves and the succession of leaves during whole-plant ontogeny. New methodological approaches, such as "process morphology" (Sattler, 1992; Jeune and Sattler, 1992), reflect the morphological flexibility of modular organization in plants, but are only semi-quantitative and cannot be directly related

to allometry (a similar approach in zoology is the "skeleton space," Thomas and Reif, 1993).

The causes of allometric variation at different levels are mutually interrelated. Static variation, which is caused by variation in ontogenetic processes that produce the structures of interest, is the "raw material" upon which natural selection can act. Response to selection, in turn, generates evolutionary changes affecting the developmental processes. One way to study these interactions is to compare empirically the patterns of variation between different levels of allometry. Various such comparisons have been made (Cheverud, 1982; Leamy and Bradley, 1982; Boag, 1984; Gibson et al., 1984; Leamy and Atchley, 1984*a*; Shea, 1985; Klingenberg and Zimmermann, 1992*a*). Most of these studies found that patterns of allometry at different levels were similar but not equal. It is not possible, however, to make further generalizations of the results because the studies differ widely in the kinds of data and methods used. As an alternative to this observational approach, the mechanisms that are the basis of allometric variation can be investigated by experimental techniques, e.g., using genetically engineered organisms (Shea et al., 1990) or directly manipulating the size of eggs or embryos (Sinervo, 1993).

Principal Components and Allometry

Under the model of simple allometry, bivariate plots of pairs of log-transformed morphometric variables are straight lines. If there are three variables, and all pairwise combinations satisfy this condition, then the data points follow a line in the three-dimensional space defined by the variables. This argument can be extended to more than three variables; data points still are arranged along a straight line under simple allometry (e.g., Teissier, 1955), but this line now is in the multidimensional space defined by all the variables. Therefore, dimensionality of morphometric variation is a prime concern of allometry (e.g., Hopkins, 1966; Sprent, 1972). As in bivariate allometry, points may be scattered around the line, rather than exactly lying on it, and one has to find a line that "optimally" fits the scatter of data points (Pearson, 1901). Jolicoeur (1963) proposed the first principal component, estimated from the covariance matrix of log-transformed measurements, as a multivariate generalization of simple allometry.

Many texts of multivariate statistics introduce principal component analysis (PCA) as a technique for summarizing most of the variation in a multivariate data set in fewer dimensions (e.g., Pimentel, 1979; Jolliffe, 1986; Flury, 1988; Flury and Riedwyl, 1988; Johnson and Wichern, 1988; Jackson, 1990; Jobson, 1992). The first principal component (PC1) is the linear combination that accounts for the maximum variance. Geometrically, it corresponds to the direction of the longest axis through the scatter of data points. Subsequent principal components take up maximal variance, subject to being orthogonal to all preceding component axes.

Figure 2-4 shows a contour ellipse of the bivariate distribution of two variables X_1 and X_2 with its centroid (mean vector) at the point labeled 0. For simplicity, data are centered by subtracting the means of X_1 and X_2, which shifts the coordinate system to the new axes x_1 and x_2, each with a sample mean of zero. Therefore, the x_1 and x_2 values themselves are the deviations from their mean; their sample variances can be calculated as the sum of squared x_1 and x_2 values divided by $(n - 1)$. In figure 2-4, one data point is labeled P, and its projection onto the x_1 axis is Q. The sum of squared x_1 values is the sum of the squared distances between 0 and Q for all data points. By the same argument, the sum of squared x_2 values corresponds to the sum of squared distances between P and Q. According to the Pythagorean theorem, the squared distance between 0 and P is the sum of the squared distances between 0 and Q and between P and Q. It follows that the sum of the squared distances of all data points from the sample centroid, divided by $(n - 1)$, is the sum of the variances of x_1 and x_2, or total variance. Now, consider the same data set after rotating the coordinate system to the directions of the principal

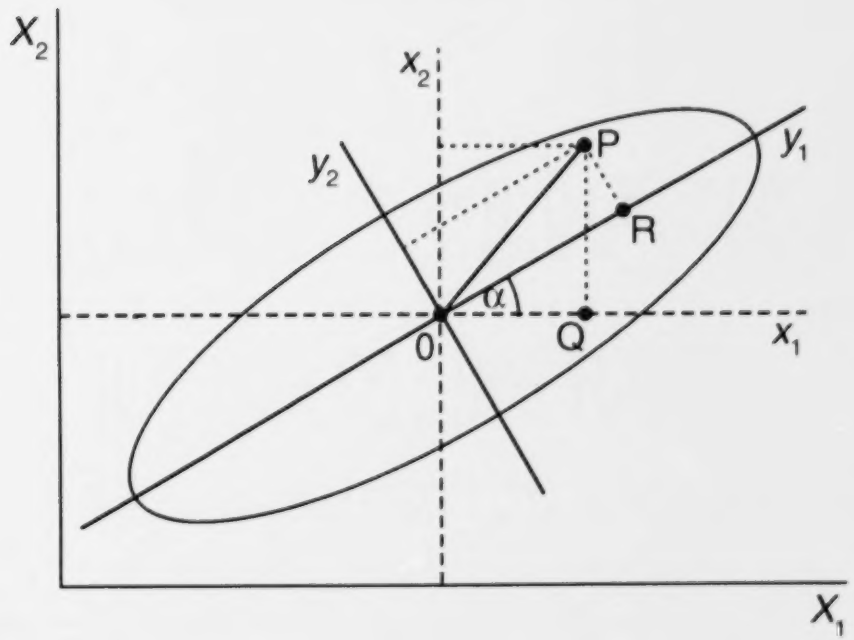


Fig. 2-4. Principal component analysis. The diagram shows the contour ellipse of a bivariate distribution with its centroid at the point labeled 0. The centered coordinates x_1 and x_2 are derived from the original variables X_1 and X_2 by subtracting their mean values. The principal components y_1 and y_2 are the directions of maximal and minimal variance, respectively. See text for details.

component axes, y_1 and y_2. Since the PC1 axis, y_1, is defined as the direction that has maximal variance, the sum of the squared distances between 0 and R, the projection of P onto the PC1 axis, is maximal. Because the rotation of the coordinate system does not change the distances between the data points and the centroid, maximizing the sum of squared distances between 0 and R also results in minimizing the sum of squared distances between P and R. The sum of squared distances between P and R, divided by $(n - 1)$, is the part of the total variance not accounted for by the PC1, i.e., the residual variance. This argument also holds for more than two dimensions: because the PC1 is the direction that has maximal variance, all other principal components taken together have minimal variance. Hence, the PC1 axis can be seen as a *least-squares fit of a straight line* to the scatter of data points in the space of log-transformed, bivariate or multivariate data. This justifies Jolicoeur's (1963) multivariate generalization of allometry (see also Hopkins, 1966).

PCA decomposes a covariance matrix **S** into eigenvectors and eigenvalues, so that $\mathbf{S} = \mathbf{BLB}'$. The matrix **B** of *eigenvectors* is used to transform the original data **X** into a set of new variables $\mathbf{Y} = \mathbf{XB}$, the principal components (PCs). The matrix **L** is the covariance matrix of the PCs, and as the PCs are uncorrelated among each other, all off-diagonal elements of **L** are zero. The diagonal elements of **L**, the *eigenvalues*, are the variances for which the associated eigenvectors account. They are difficult to interpret by themselves, because they depend on the measurement units and the base of the logarithm used to transform the data. However, the proportion of the total variance for which the PC1 accounts is important to assess how well the model of simple allometry fits the data.

PCA can be interpreted geometrically as a rotation of the coordinate system. The PC axes are aligned with the directions of the axes of the multidimensional "scatter ellipsoid" (in two dimensions, this is an ellipse, Fig. 2-4). The PC coefficients of the original variables can be interpreted as "direction cosines," i.e., the cosine of the angle between the PC axis and the coordinate axis of the respective variable (α for PC1 and x_1 in Fig. 2-4). PC axes are mutually orthogonal, and the vectors of PC coefficients are usually normalized to have unit length, i.e., so that the squares of the coefficients sum up to unity ($\mathbf{b}'\mathbf{b} = 1$, where **b** is an eigenvector). As a result, the coefficient values depend on the number of variables. Nevertheless, translating PC coefficients to bivariate allometric coefficients (Huxley's α) is quite easy (Jolicoeur, 1963; Shea, 1985). The ratio of PC1 coefficients for two variables corresponds to their bivariate allometric coefficient. For example, in a study of two species of voles, Airoldi and Flury (1988) found that the PC1 coefficients of skull length, width, and height were approximately 2/3, 2/3, and 1/3, respectively. Thus, in allometric plots using skull length as the independent variable, skull width would be isometric with a slope of about 1, whereas skull height, with a slope of about 0.5, would show strong negative allometry. With p variables, isometry in all pairwise combinations of variables results in a PC1 in which all coefficients are equal, and have the value $1/\sqrt{p}$. Isometry can be assessed with Anderson's (1963) test, which is based on normal theory (Pimentel, 1979: 70; Flury, 1988: 34), or by comparison with jackknifed or bootstrapped confidence intervals (see below; e.g., Diaconis and Efron, 1983; Gibson et al., 1984; Klin-

genberg and Zimmermann, 1992*a*). Multiplying PC1 coefficients by $\sqrt{p}$ yields values that can be interpreted as bivariate allometric coefficients for each of the variables against a measure of "overall size" (a weighted geometric mean of all variables). These allometric coefficients or the principal component coefficients can be graphed as Huxley's (1932) growth gradients (e.g., Boitard et al., 1982; Solignac et al., 1990), or they can be displayed on diagrams of the measurements, such as the truss network (Strauss and Bookstein, 1982; Bookstein et al., 1985). Another type of graphical display for PC coefficients is the biplot (e.g., Marcus, 1993).

Empirical comparisons of bivariate and multivariate approaches found that both estimated corresponding patterns of allometry (Davies and Brown, 1972; Shea, 1985). Jungers and German (1981) criticized the multivariate approach because allometric coefficients derived from principal components of skeletal measurements did not match those from bivariate regressions on a "known" variable for size that was not included in the analysis, either body weight or length. Hills (1982) showed that these discrepancies disappear if one considers allometry between the traits and the size measure that is taken as a reference, e.g., between "skeletal size" and body weight.

It is possible to perform PCA using a correlation matrix instead of a covariance matrix (see also Pimentel, 1979; Bookstein et al., 1985; Johnson and Wichern, 1988). This corresponds to an analysis of standardized variables. Geometrically, it means that all the variables are adjusted to have standard deviations of 1 by stretching or shrinking their coordinate axes before the analysis. This maneuver can be used to remove scaling effects if variables are measured in different units. Standardizing may also be useful if one is only interested in ordination of specimens, as in some applications in systematics, where giving equal weight to all variables may be more important than scaling. For allometry, however, scaling is essential. After removing scale by standardization one still can determine whether the data points lie along a straight line, but it is impossible to estimate allometric coefficients, because standardization changes the direction of the allometric axis. A simple hypothetical example, constructed from purely allometric variation without any residual scatter, can show this. Let the multivariate allometric coefficients (eigenvector) be 2/3, 2/3, and 1/3; these coefficients correspond to strong deviations from isometry. Then, the covariance matrix is a multiple of

$$\begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix} \text{ and the correlation matrix is } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

The PC1 of this correlation matrix has coefficients that all take the value $1/\sqrt{3}$, and thus falsely indicate "isometry", with no residual variation at all (see also Johnson and Wichern, 1988: 350). Likewise, in allometric interpretations of PCAs using correlation matrices of real data (e.g., Teissier, 1955; Somers, 1986), all "non-isometric" variation that may be inferred from PC1 coefficients merely results from the residual scatter about an allometric relationship, but does not reflect allometry. *Therefore, it is crucial to use the covariance matrix for allometry.* Routines for PCA in many statistical software packages use the correlation matrix as the default option. Users of these pro-

grams should make sure to specify the option for PCA using the covariance matrix.

A related point is the transformation of data to logarithms. There are numerous practical and theoretical reasons why it is often useful to transform data to logarithms (Pimentel, 1979; Reyment et al., 1984; Bookstein et al., 1985; Bookstein, 1991; Reyment, 1991). For ontogenetic allometry, Huxley (1932) justified the use power functions, and therefore also of logarithms, by his rule of constant ratios among specific growth rates of different organs (see also Reeve and Huxley, 1945; Günther, 1975; Katz, 1980; Bookstein et al., 1985; Shea, 1985; Blackstone, 1987). Such a theoretical justification, however, is more difficult to find for static and evolutionary allometry. The multiplicative nature of growth processes also may be important for these levels, because all variation in morphological structures is due to variation in the developmental processes that generate them. Mosimann (1970) and Mosimann and James (1979) pointed out statistical advantages of the log-normal distribution (but see Smith, 1993, for biases in predicted values from allometric regression). In practice, log-transformation often renders relations among variables more linear and also can make variances more homogeneous. Finally, log-transformed data are independent of measurement units (e.g., millimeters or the units of an eyepiece micrometer), but retain the information about scale (e.g., lengths vs. surfaces). It does not matter which base for the logarithms is chosen, as long as the same base is used consistently throughout a given analysis.

Results of PCAs are estimates of allometric patterns in the populations from which the study samples are drawn. To assess how reliable these estimates are, standard errors or confidence intervals should be calculated. Formulas for these statistics (e.g., Flury and Riedwyl, 1988) are based on the assumption of multivariate normal distribution and on large sample sizes. In most allometric studies, however, the distribution of data cannot be assumed to be multivariate normal. In ontogenetic allometry, for example, the distribution of measurements depends on the age composition in the sample, as well as on the growth dynamics of the structures investigated. In these cases, the bootstrap and jackknife procedures are helpful (an excellent introduction is Efron and Tibshirani, 1993; other useful references are Diaconis and Efron, 1983; Efron and Gong, 1983; Efron and Tibshirani, 1986; Manly, 1991). The bootstrap is a computer-intensive procedure that substitutes repeated sampling from the sample distribution for a theoretical model of that distribution. The only assumption that must be made is that the specimens have been sampled randomly. Applications of the bootstrap to PCA include Diaconis and Efron (1983), Stauffer et al. (1985), Daudin et al. (1988), and Efron and Tibshirani (1993). In multivariate allometry, Gibson et al. (1984) and McGillivray (1985) used the jackknife, whereas Klingenberg and Froese (1991) and Klingenberg and Zimmermann (1992*a*, *b*) used the bootstrap. Marcus (1990) compared the jackknife and the bootstrap with each other and with results from large-sample theory.

The fundamental idea of the *bootstrap*, and the procedures to apply it, follow immediately from the definitions of standard errors and confidence intervals for a statistic θ (e.g., mean value or PC coefficients) estimated from a sample of n specimens. Both standard errors and confidence intervals provide answers to the same question: If the same study were repeated numerous times, estimating θ from a sample with n specimens each time, how variable would

the estimates be? The standard error of a statistic is the standard deviation of these estimates, and a confidence interval is the interval containing a certain percentage of the estimates (e.g., the 95% confidence interval is delimited by the 2.5% and 97.5% quantiles). This is exactly what the bootstrap does, assuming that the sample distribution is representative of the totality of organisms about which statements are made (e.g., all members of a local population, all females of a species). Repeatedly, a "bootstrap sample" of n specimens is drawn randomly, with replacement, from the original sample, and θ is estimated for each bootstrap sample. The standard deviation of these estimates is the bootstrapped standard error, and confidence intervals can be derived from the distribution of bootstrap estimates (for details, see Efron and Tibshirani, 1993). About 100 bootstrap samples usually are sufficient to establish standard errors, but at least about 1000 are necessary for confidence intervals (Efron and Tibshirani, 1993). The bootstrap can even be used for hypothesis tests, by generating bootstrap replicates of a test statistic that under a particular null hypothesis, and then comparing the resulting distribution to the test statistic calculated for the observed data (Efron and Tibshirani, 1993; for a related topic, permutation tests, see also Manly, 1991). An advantage of the bootstrap is that it can be adapted to the particular design of a study. For instance, if there are discrete growth stages that can be identified unambiguously (see Fig. 2-1), there may be no sampling error in the stage composition of the data set (e.g., a design with equal numbers from each stage). The bootstrap procedure for ontogenetic allometry can be adapted by drawing "bootstrap subsamples" from these stages separately. These are then pooled into one bootstrap sample and principal components are calculated (Klingenberg and Zimmermann, 1992*a*).

Analyses of Multiple Groups

Many morphometric studies deal with several groups of specimens, e.g., sexes, different species or ecomorphs. In all these cases, variation within and between groups has to be separated. Otherwise, levels of allometry may be confounded, or within-group variability may invalidate discrimination between groups. Separation of size-related variation within groups from between-group differences has been a traditional topic in morphometrics (Burnaby, 1966; Gower, 1976; Reyment and Banfield, 1976; Pimentel, 1979; Humphries et al., 1981; Thorpe, 1983; Reyment et al., 1984; Bookstein et al., 1985; Rohlf and Bookstein, 1987; Marcus, 1990; Reyment, 1991).

Multivariate comparisons of allometric patterns often focus on the directions of the major axes of scatter ellipsoids in several groups. A straightforward measure for differences between two groups is the angle between their first principal components. For normalized principal components (i.e., squared coefficients sum up to unity), the angle α between components $\mathbf{b}$ and $\mathbf{c}$ in two groups is the arccosine of the inner product of the two vectors, $\alpha = \arccos\left(\sum b_i c_i\right) = \arccos(\mathbf{b}'\mathbf{c})$ ($\mathbf{b}'\mathbf{c}$ is sometimes called vector correlation; note that this is *not* the correlation between corresponding elements of the two vectors; see also Pimentel, 1979; Bryant, 1984). Angles can even be calculated from published tables of PC coefficients. Applications of angular comparisons include Boitard et al. (1982), and Gibson et al. (1984). Cheverud (1982) and

Klingenberg and Zimmermann (1992*a*) used Monte Carlo simulations of angles between random vectors to assess statistical significance. For the example in this paper, I used the bootstrap to test the more appropriate null hypothesis of equal PC vectors (0° angles).

Another method for comparison among multiple groups of organisms is based on a multivariate ordination of the directions of allometric axes (Solignac et al., 1990; Klingenberg and Froese, 1991; Klingenberg and Spence, 1993). The first principal component of each group is considered as a "data point" in the space spanned by the coefficients of the original variables. The vectors of PC1 coefficients are entered as observations in an ordination by a second PCA. The results of this analysis can then be displayed as plots of the "meta-PC" scores. Bootstrap estimates of allometric coefficients can be used to draw confidence ellipses (e.g., Owen and Chmielewski, 1985; Johnson and Wichern, 1988) as a visual indication of statistical accuracy (Klingenberg and Froese, 1991; Klingenberg and Spence, 1993).

Comparisons of allometry within several groups often, but not always, show that the coefficients of the PC1s differ only minimally. In these cases, it may be feasible to use one of the models of common covariance structure (Fig. 2-5; Airoldi and Flury, 1988; Flury, 1988). These models are based on the assumption that the groups share a common allometric pattern, i.e., that the major axes of their scatter ellipsoids are parallel. Therefore, the differences between the observed PCs of the samples are regarded as effects of sampling error. PCA, however, is a procedure for analyzing variation in a single sample, and therefore the method needs to be generalized for the context of multiple groups. Usually, the PC1 of the pooled within-group covariance matrix has been used to characterize this common allometric pattern, e.g., in multigroup PCA (Pimentel, 1979; Thorpe, 1983), Burnaby's procedure (Burnaby, 1966; Rohlf and Bookstein, 1987), and in the shearing procedure (Humphries et al., 1981; Bookstein et al., 1985; Rohlf and Bookstein, 1987). Airoldi and Flury (1988)

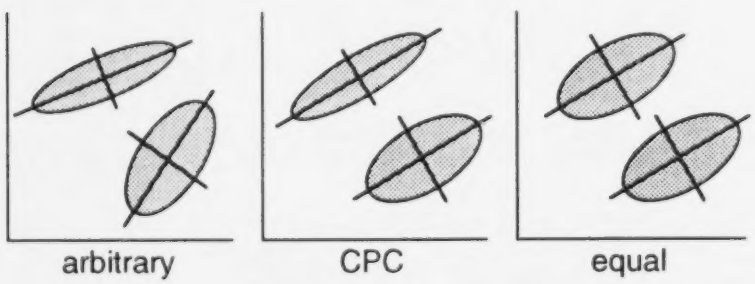


Fig. 2-5. Three levels of similarity between covariance structures. Groups can have arbitrarily different covariance matrices; scatter ellipses then differ both in the directions and lengths of their principal axes. Under the common principal component model (CPC), groups share the same directions of principal axes, but may differ in the amount of variation associated with each axis. Groups with equal covariance matrices have the same directions and lengths of principal axes. More details, including additional levels of similarity, are given by Airoldi and Flury (1988) and Flury (1988).

criticized the use of the pooled within-group covariance matrix, because it implicitly assumes that the covariance matrices of all groups are identical (Fig. 2-5, right panel). They proposed an alternative procedure, common principal component analysis (CPCA), which only assumes that the PCs are common to all groups (see also Flury, 1988; Flury and Riedwyl, 1988). Whereas the directions of the principal axes are assumed to be the same, the amount of variation associated with each PC can vary between groups (Fig. 2-5, center). CPCA is available in the NTSYS software package, FORTRAN routines are contained in the IMSL/STAT library (routines KPRIN and DKPRIN), a MATLAB program was written by L. F. Marcus, and a SAS/IML version is available from the author. In applications of CPCA to multivariate allometry, the first common principal component (CPC1) is interpreted as an allometric pattern shared by all groups (Airoldi and Flury, 1988; Klingenberg and Zimmermann, 1992*a*, *b*; Klingenberg and Spence, 1993).

Discrimination between groups is often difficult because of allometric variation within groups. Especially in organisms with indeterminate growth, the amount of within-group variation may far exceed between-group differences. For instance, fish can increase in size by several orders of magnitude during their life cycle. Other causes, such as nutrition, also contribute to variability within groups (Bernays, 1986; Patton and Brylski, 1987; Meyer, 1990; Smith and Palmer, 1994). Depending on the particular organisms of interest, within-group variation is mostly ontogenetic or mostly static allometry, or a mixture of both. Because most of this variation is often confined to a single dimension, along the allometric axis, it can be removed from an analysis by eliminating the variation in that direction. This approach uses allometry as a *criterion of subtraction* (Gould, 1975), as it has been done traditionally in bivariate studies. The central assumption of all methods for "size correction," is that the groups share the same allometric vector. If the groups have different allometric patterns, "size correction" is not possible, because all corrections that are suitable for one group will not work in other groups. Several methods have been developed following this principle (Burnaby, 1966; Gower, 1976; Humphries et al., 1981; Thorpe, 1983; Bookstein et al., 1985; Rohlf and Bookstein, 1987). None of these methods, however, can be a substitute for careful examination of the specimens: hidden heterogeneity within the groups, e.g., undetected sex dimorphism or cryptic species, may invalidate the entire analysis.

The procedure proposed by Burnaby (1966) eliminates the effects of growth from multivariate data by projecting data points onto a subspace that is orthogonal to the growth vector (see also Gower, 1976; Rohlf and Bookstein, 1987; Reyment, 1991). This *growth-invariant* subspace has one dimension fewer than the original space. With three variables, for example, the growth-invariant space is a plane, and for two variables, it is a line (Fig. 2-6). The growth-adjusted data are coordinates of the projected points, expressed in the coordinate system of the original variables. The growth-invariant data for an $n \times p$ data matrix $\mathbf{X}$ and a $p \times 1$ growth vector $\mathbf{b}$ can be obtained as $\mathbf{X}\left(\mathbf{I} - \mathbf{b}(\mathbf{b}'\mathbf{b})^{-1}\mathbf{b}'\right)$, where $\mathbf{I}$ is an identity matrix of rank p. With a normalized vector, such as a principal component, the formula simplifies to $\mathbf{X}(\mathbf{I} - \mathbf{b}\mathbf{b}')$. Usually, the PC1 of the pooled within-groups covariance matrix has been used as the growth vector in Burnaby's procedure (e.g., Reyment and Banfield, 1976; Riska, 1981; Rohlf

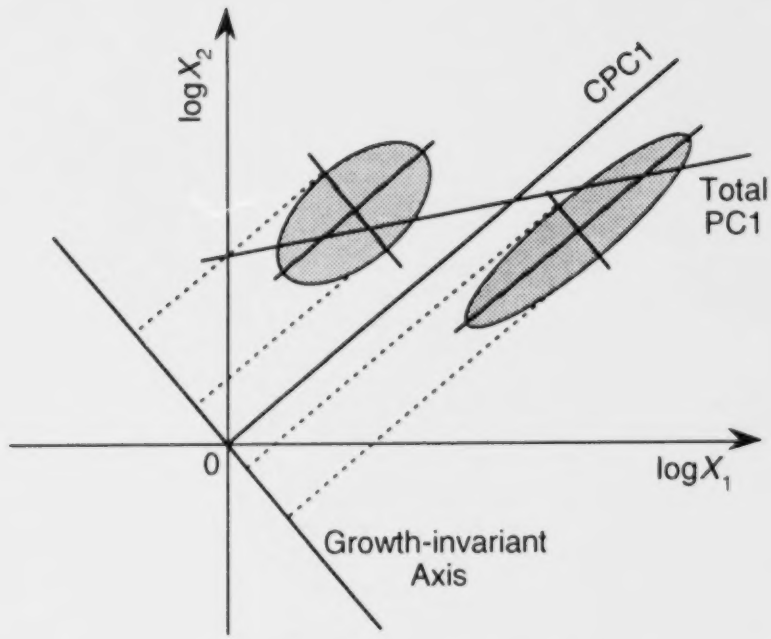


Fig. 2-6. Burnaby's procedure in two dimensions. The model of common principal components is appropriate, because the principal axes of the scatter ellipses in the two groups are parallel. Therefore, the groups share a common axis of allometric growth, the CPC1. All the variation is projected onto an axis perpendicular to the CPC1 by setting the CPC1 scores to zero. Subsequent analyses focus on between-group differences in this "growth-invariant" axis. The PC1 of the combined samples ("Total PC1") confounds variation within and between groups.

and Bookstein, 1987). Because CPCA is based on less stringent assumptions (see above), the CPC1 seems preferable as an estimate of a common allometric pattern. This version of Burnaby's technique is equivalent to a procedure involving three consecutive steps: (1) a rigid rotation to the common principal components, (2) setting the CPC1 scores of each data point to zero, and (3) rigid rotation back to the original coordinate system.

For analyses of growth-adjusted data, it may be more convenient to omit step (3). First, common principal components are computed as an estimate of within-group variation, and the CPC scores, except for the CPC1, are used as variables in subsequent analyses. The results of these analyses, e.g., discriminant analysis or MANOVA, are identical to the results based on data adjusted by Burnaby's original procedure, but within-group covariance matrices are of full rank. This technique is almost identical to the one proposed by Thorpe (1983; with the PC1 of the pooled within-groups covariance matrix as an estimate of the allometric axis), which has been used in numerous studies (e.g., Wiig, 1985; Thorpe and Baez, 1987; Corti et al., 1988; Lessa and Patton, 1989).

Interpreting the values resulting from Burnaby's procedure is somewhat difficult. With some caution, they can be seen as measurements adjusted to "unit size". This is possible for most morphometric data sets, because the CPC1 has high positive correlations with all variables, and therefore can be interpreted as an overall size variable. Setting the CPC1 score for log-transformed measurements to zero corresponds to setting the untransformed value of overall size to one. The adjusted variables therefore allow to compare organisms of different sizes by rescaling all measurements allometrically to unit size. For example, in figure 2-6 the group to the left has higher X_2 and lower X_1 values at corresponding sizes than the group to the right. The difference between this approach and comparisons of ratios or geometric shape is that Burnaby's procedure takes into account the shape changes caused by allometric growth.

Another way to understand Burnaby's procedure is by analogy to the most familiar method of size correction in bivariate allometry, regression residuals. Regression residuals are an appropriate way of correcting for size if one is interested in the relationship of a dependent variable y (e.g., an ecological or physiological parameter) to a size variable x known *a priori* and measured independently (e.g. body weight). Then, the residuals are the deviations of actual measurements from the value expected for an "average" specimen of that particular size. These deviations are computed in the direction of the y axis, by subtracting the y value estimated by regression from the observed value. Hence, the interpretation of allometry as a "criterion of subtraction" (Gould, 1975) can be taken literally. In multivariate allometry, however, there is no *a priori* size variable that can be measured independently from the other variables. All measurements are affected simultaneously by overall size, which only can be estimated from them. As I explained above, the PC1 is a good choice for such an estimate. The "residuals" are subsequent PCs, which are perpendicular to the PC1 [step (2) above]. In bivariate regression, comparison of intercepts of several groups makes sense only when the slopes in all groups are equal. Analogously, Burnaby's procedure only works if all the groups share a common allometric pattern, as it is estimated by the CPC1. (For groups that differ in their growth vectors, Burnaby [1966] suggested removing all growth vectors from the data; after this, however, there may not be much meaningful variation left to study [e.g., Humphries et al., 1981].)

Because it includes only rigid rotations, Burnaby's procedure conserves the spatial relationships among data points in all directions except that of the growth vector. Therefore, data adjusted by this technique can be used to quantify variation perpendicular to the growth trajectories. For instance, lateral transpositions of growth trajectories (also called vertical transpositions) can be distinguished from group differences produced by shifts along the growth axis (ontogenetic scaling, Gould, 1975; Shea 1985, 1992). These two types of group differences, as well as within-group variation, are confounded in PCA of combined samples (Fig. 2-6; for discussion, see Voss et al., 1990; Voss and Marcus, 1992; Klingenberg and Spence, 1993). Burnaby-adjusted data also can be used to examine within-group deviations from simple ontogenetic allometry, such as curvatures of the growth trajectories. Klingenberg and Spence (1993) used a MANOVA of Burnaby-adjusted data to separate lateral transposition from non-allometric growth in a data set containing samples of six

discrete ontogenetic stages from each of six waterstrider species (i.e., a data structure corresponding to Fig. 2-1). The PC1 scores of the between-species matrix can be used to display lateral transposition, and likewise the between-stage matrix for non-allometric growth.

Procedures for "size correction" in several groups all assume that the groups share the same allometric pattern. But what if that is not true? In this case, the formal procedures fail, and one has to find a visual and at best "semi-statistical" way to assess group differences. A very useful technique of this kind is the "tomographic representation" introduced by Boitard et al. (1982). First, they calculated the PCs for the covariance matrix of the pooled data. Then, they plotted the second versus third PC scores separately for data points grouped by their PC1 scores. Finally, they combined these layers into a plot showing the scatter ellipsoids "suspended" by their within-group PC1 axes in a "box" representing the space of the first three PCs of the pooled data.

EXAMPLE: GEOGRAPHIC VARIATION

Some of the techniques described above are applied in a simple example, a morphometric data set on geographic variation in the waterstrider *Gerris costae*, taken from a larger study (Klingenberg, 1992). Only adult specimens were measured, and because these bugs do not grow after they reach the adult stage, I made no attempt to correct for size in the original study. Here, I reanalyze a part of the data using allometric techniques.

Three samples, each representing a different subspecies, are included here: the nominate subspecies *G. c. costae* is represented by a sample from the Swiss Alps ($N = 32$), *G. c. fieberi* is represented by a sample from northern Greece ($N = 33$), and *G. c. poissoni* by a sample from the eastern Pyrenees in France ($N = 28$). In this example, I only consider adult males. Allometric variation within the three samples is therefore purely static allometry.

Four measurements were chosen for the example: (1) total thorax length, (2) the length of the first antennal segment, and (3) the middle femur length and (4) hind femur length. The raw data are presented in the Appendix. Data were transformed to natural logarithms. For the convenience of presentation, variances and covariances are multiplied by 10^4 (this also applies to eigenvalues). The covariance matrices are

$$\begin{bmatrix} 12.67 & 8.39 & 8.75 & 11.09 \\ 8.39 & 14.34 & 9.04 & 11.61 \\ 8.75 & 9.04 & 12.05 & 13.63 \\ 11.09 & 11.61 & 13.63 & 17.83 \end{bmatrix}$$ for the sample from the Alps,

$$\begin{bmatrix} 5.55 & 2.07 & 2.53 & 4.06 \\ 2.07 & 9.66 & 4.03 & 4.67 \\ 2.53 & 4.03 & 7.82 & 9.02 \\ 4.06 & 4.67 & 9.02 & 12.98 \end{bmatrix}$$ for the sample from Greece, and

$$\begin{bmatrix} 4.93 & 5.97 & 3.90 & 5.16 \\ 5.97 & 18.65 & 6.78 & 8.35 \\ 3.90 & 6.78 & 6.94 & 7.56 \\ 5.16 & 8.35 & 7.56 & 10.70 \end{bmatrix}$$ for the sample from the Pyrenees.

As in many morphometric data sets, all covariances are positive and most of them are relatively high. There are, however, some differences between groups and among variables. The sample from the Alps is more variable than the other two samples. While the two femur lengths are highly correlated in all groups, correlations involving thorax length and the first antennal segment tend to be lower and more variable (e.g., the correlation between these two measurements is only 0.28 in the Greek sample).

Principal components were computed for each sample, and parametric standard errors for the estimates of PC coefficients and eigenvalues were calculated using the formulas given by Flury (1988). Moreover, standard errors were also determined with a bootstrap procedure. For each group, 1000 bootstrap samples were randomly drawn (with replacement), and principal components were computed for each. (This number of replications is more than actually needed for standard errors; Efron and Tibshirani, 1993.) The standard deviations of these 1000 estimates of the PC coefficients or eigenvalues are their bootstrap standard errors.

The amount of variance (± parametric and bootstrapped standard errors), for which PC1 accounts, is 46.1 (± 11.7, 12.5) or 81% of total variance in the sample from the Alps, for 24.1 (± 6.0, 4.5) or 67% in the Greek sample, and for 31.4 (± 8.5, 7.2) or 76% in the sample from the Pyrenees. Standard errors determined by the two approaches are similar in magnitude, although the bootstrap standard error for the Greek sample differs from the parametric estimate by about 25%. The percentages of total variance taken up by the PC1 are quite typical for static allometry (e.g. Cuzin-Roudy, 1975; Klingenberg and Zimmermann, 1992*a*). The estimated coefficients of the PC1 and their parametric and bootstrapped standard errors are

$$\begin{bmatrix}0.441\\0.471\\0.477\\0.597\end{bmatrix}, \begin{bmatrix}0.057\\0.061\\0.036\\0.034\end{bmatrix}, \text{and} \begin{bmatrix}0.060\\0.055\\0.032\\0.030\end{bmatrix} \text{ for the sample from the Alps,}$$

$$\begin{bmatrix}0.269\\0.409\\0.527\\0.695\end{bmatrix}, \begin{bmatrix}0.085\\0.119\\0.045\\0.054\end{bmatrix}, \text{and} \begin{bmatrix}0.093\\0.113\\0.045\\0.060\end{bmatrix} \text{ for the Greek sample, and}$$

$$\begin{bmatrix}0.315\\0.695\\0.400\\0.507\end{bmatrix}, \begin{bmatrix}0.050\\0.082\\0.057\\0.070\end{bmatrix}, \text{and} \begin{bmatrix}0.037\\0.077\\0.057\\0.078\end{bmatrix} \text{ for the sample from the Pyrenees.}$$

The estimates of PC coefficients are fairly stable, as indicated by the relatively small standard errors. Bootstrap standard errors agree well with parametric estimates. The coefficient values clearly do not conform to the pattern for overall isometry, where all coefficients would be equal to $1/\sqrt{p}$, i.e., the vector [0.5, 0.5, 0.5, 0.5]'. This is confirmed by Anderson's test (Anderson, 1963; Pimentel, 1979: 70; Flury, 1988: 34), which was significant for all three groups (Alps: $\chi^2 = 11.52$, $P < 0.01$; Greece: $\chi^2 = 20.37$, $P < 0.001$; Pyrenees $\chi^2 = 23.82$, $P < 0.0001$; df = 3 in each case).

Coefficients vary considerably between samples. In the sample from the Alps, the middle femur (third measurement) is almost isometric relative to thorax length (first measurement), as the ratio of their coefficients is 0.477 / 0.441 = 1.08. In the Greek sample, however, the corresponding ratio is 0.527 / 0.269 = 1.96, which indicates strong positive allometry of the middle femur relative to the thorax. While the lengths of the first antennal segment and of the middle femur (second and third measurements) are almost isometric in the sample from the Alps, the middle femur shows positive allometry relative to the first antennal segment in the Greek sample, and negative allometry in the Pyrenees. As an overall measure of these differences, I computed the angles between PC1 axes of different groups; they are 12.22° between Alps and Greece, 16.29° between Alps and Pyrenees, and 21.19° between Greece and the Pyrenees.

If we expect a common allometric pattern, these angles seem quite large. But are the differences from zero statistically significant? Because there is no parametric test for the angles between PCs, I used the bootstrap approach to test the null hypothesis that the common principal component model holds. The bootstrap test procedure is straightforward: use the data to generate a modified data set that conforms to the null hypothesis, then repeatedly draw bootstrap samples and compute the test statistic, and finally compare the empirical distribution of the test statistic to the value calculated for the original data (Efron and Tibshirani, 1993). To produce a data set conforming to the CPC model, I rotated the data in each sample so that the within-group PC axes are aligned exactly with the CPC axes. This is easy to do because the matrix of PC coefficients $\mathbf{B}$ can be used to rotate the data points from the original coordinate system to the PC coordinates $\mathbf{Y} = \mathbf{XB}$, whereas the transpose of $\mathbf{B}$ performs the reverse rotation $\mathbf{YB}' = \mathbf{X}$. A data set conforming to the CPC model can be obtained by using each group's own PC coefficients for the first rotation, but the coefficient matrix from a CPCA for the reverse rotation. Thus, the modified data in the i-th group are $\mathbf{X}_i^* = \mathbf{X}_i\mathbf{B}_i\mathbf{B}_{CPC}'$, where $\mathbf{B}_i$ is the matrix of within-group PC coefficients, and $\mathbf{B}_{CPC}$ is the matrix of CPC coefficients (using the PC scores for each group, $\mathbf{Y}_i = \mathbf{X}_i\mathbf{B}_i$, is equivalent). For this test, 5000 bootstrap replications were performed. For each replication, bootstrap samples were drawn from the modified data of all three groups, and the angles between the PC1 axes of the three groups were computed. The 95% quantiles of the angles are 18.56° between the samples from Alps and from Greece, 16.32° between the Alps and the Pyrenees, and 20.59° between the Pyrenean and the Greek samples. Therefore, the two angles that involve the sample from the Pyrenees seem to indicate a significant difference from the Greek sample, and borderline significance for the difference from the Alps. Because there are three comparisons, however, the chance of rejecting a true null hypothesis at the $\alpha = 0.05$ level by chance is $1 - (1 - \alpha)^3 = 1 - 0.86 = 0.14$. The Bonferroni technique can be used to adjust the significance level for m individual comparisons to a new significance threshold at $(1 - \alpha/m) = 98.33\%$. None of the comparisons was significant at this adjusted level (98.33% quantiles are 22.23, 20.63, and 25.48). Therefore, the data seem to be consistent with a model of an allometric pattern that all three groups have in common.

This preliminary conclusion justifies using common principal components to estimate an allometric pattern for all three groups simultaneously. Again,

standard errors were estimated both by using formulas based on large sample theory (Flury, 1988) and by bootstrapping with 1000 bootstrap iterations. The estimates of the CPC1 coefficients with their parametric and bootstrapped standard errors are

$$\begin{bmatrix}0.361\\0.510\\0.482\\0.614\end{bmatrix}, \begin{bmatrix}0.034\\0.047\\0.024\\0.028\end{bmatrix}, \text{ and } \begin{bmatrix}0.039\\0.053\\0.024\\0.031\end{bmatrix}, \text{ respectively.}$$

Although in a CPC model there is only one set of eigenvectors, which is shared by the groups, each group has its own eigenvalues. The variance taken up by the CPC1 (± parametric and bootstrap standard errors) in each sample is 45.8 (± 11.4, 12.2) for the Alps, 23.6 (± 5.8, 4.7) for Greece, and 30.0 (± 8.0, 7.4) for the Pyrenees. This corresponds to 80.4%, 65.7%, and 72.8% of the total variance in the respective samples. These values are only a little lower than corresponding values from the separate one-group PCAs of each sample; one common allometric pattern can account for almost as much of the variation as the PC1 of each group separately. Further support for a common model comes from the log-likelihood ratio test of the CPC model (Flury, 1988; Airoldi and Flury, 1988), which does not show a significant difference from the unrestricted model, in which every group has its own PCs ($\chi^2 = 10.55$, df = 12, $P = 0.57$). For the same χ^2 statistic, the bootstrap test with 5000 iterations yielded a 95% quantile of 20.90, which agrees nicely with the corresponding value of 21.03 in statistical tables (the same bootstrap replicates as for the angles between one-group PCs, above). The angles between the CPC1 axis and the one-group PC1 axes indicate that the CPC estimate is a "compromise" between the one-group PCs (5.2° for Alps; 9.4° for Greece; 13.5° for Pyrenees). In the bootstrap test, none of these angles exceeded the 98.33% quantiles (Bonferroni adjustment for $\alpha = 0.05$; quantiles are 9.85°, 18.25°, and 17.75°, respectively). Therefore, it is reasonable to assume that the three groups share the same allometric pattern, and that the differences between the PC1 estimates of individual groups are due to sampling error.

To examine whether the differences between groups are simply extensions of within-group variation, I used Burnaby's approach. "Size-invariant" variation between groups was analyzed by a MANOVA of CPC scores, omitting the CPC1. The first eigenvector of the resulting between-groups matrix indicates the axis that contains the most variation among groups, subject to being perpendicular to the within-group allometric axis. Because there are three groups, the matrix of between-groups sums of squares (mean squares would yield equivalent results) has only two non-zero eigenvalues. The first eigenvector of this matrix accounted for 82% of the total between-group variation, and it therefore summarizes most of the differences between samples after adjusting for "size."

Figure 2-7 is a plot of this axis of group differences versus the CPC1 scores. The Greek specimens differ from the Alpine ones mainly by their higher CPC1 scores, which indicate greater overall size. Therefore, the bugs of the Greek sample can be seen as "scaled-up" versions of their counterparts from the Alps, corresponding to intraspecific scaling along the axis of static allometry (Gould, 1975; although these data deal with subspecies, and not with different species,

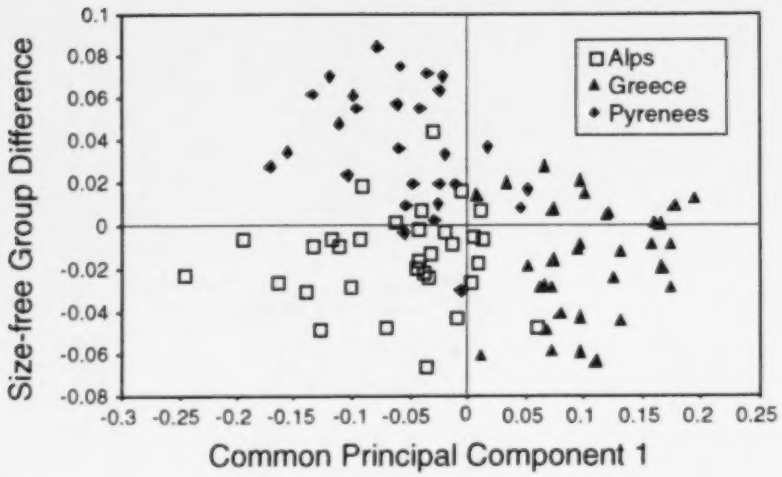


Fig. 2-7. Geographic variation in the waterstrider *Gerris costae*. The first common principal component of the three samples (abscissa) can be interpreted as a measure of overall size. It indicates that Greek specimens are considerably larger than those from the Alps or the Pyrenees. The vertical axis is the PC1 of the between-group matrix from a MANOVA of "size-free" data (the scores of CPC2–CPC4), and summarizes differences between groups independent of within-group allometry. It shows that the samples from the Alps and Pyrenees are separated fairly well, although they are of similar overall size.

I do not think it would be helpful to coin a new term). The differences between the samples from the Alps and Pyrenees are largely unrelated to within-group variation. These conclusions are also supported in a more quantitative way by the Mahalanobis distances between the three groups. The Mahalanobis D^2 value, computed from the second to fourth CPC scores, is clearly smaller between the samples from the Alps and from Greece ($D^2 = 2.38$) than between either of these and the sample from the Pyrenees (Alps vs. Pyrenees $D^2 = 6.55$; Greece vs. Pyrenees $D^2 = 4.92$). These values, however, also show that scaling does not account for all the difference between the specimens from the Alps and from Greece.

The bootstrap technique has several advantages for assessing statistical accuracy and even for hypothesis testing in morphometric analyses. First, it does not require that the data conform to any particular probability distribution as other techniques do. Nevertheless, assumptions about distributions can be incorporated in the simulations (parametric bootstrap; Efron and Tibshirani, 1993). For the present example, the parametric bootstrap for a multivariate normal CPC model gave results similar to the nonparametric results presented above. Second, the bootstrap can be used for any test statistic, even if its statistical properties are unknown. In the example, I extensively used the angles between PC axes, because angles are particularly intuitive as a measure of overall similarity for allometric vectors. Moreover, the bootstrap can be adapted easily to a variety of experimental designs or hypothesis tests. For the

bootstrap test of the CPC model, rotation of the original data was sufficient to generate a data set conforming to the null model. With tools such as the singular value decomposition (e.g., Marcus, 1993), a variety of other null models could be simulated with real data.

These advantages have a cost, however, as the bootstrap technique is based on massive amounts of numerical calculation. As computers become faster and cheaper, this may not be a serious problem except for extremely large or complex data sets. For this example, all bootstrap analyses were done by a personal computer with a 486/50MHz processor, using SAS/IML software (version 6.08). Although the number of bootstrap replications was substantial, the computation time was moderate. For standard errors in one-group PCA, with 3000 analyses (1000 bootstrap replications for three groups), the entire bootstrap procedure took less than 1.5 minutes. The 1000 bootstrap replicates for CPCA took about 45 minutes, much longer than for ordinary PCA, because the computational procedure for CPCA is more complex. The most effort was required for the bootstrap test: the 5000 iterations, each with three one-group PCAs and CPCA, took a little more than three hours. The high number of bootstrap replications for the test was necessary because one is interested in the tails of the test statistic's empirical distribution (Efron and Tibshirani, 1993). The computational effort used for this example shows that the bootstrap is a reasonable option, even with a personal computer and for relatively complex problems.

Alternative Approaches in Allometry

In the preceding sections, I presented allometry in a pragmatic way, extending the familiar logic of two-dimensional scatter plots to a multivariate context. This approach focuses on the patterns of variation by determining amount, dimensionality, and direction of morphometric variation in the space of log-transformed variables.

Some readers may have noticed that I only used the words "size" and "shape" in a rather informal way, although they are the central concepts for other approaches in morphometrics (Bookstein et al., 1985; Bookstein, 1989, 1991, 1993; Rohlf, 1990; Rohlf and Marcus, 1993). In studies using the approach described above, "size" and "shape" may appear in interpretations of the results, but they are not parts of the analyses themselves. The analyses are exploratory or they test simple hypotheses about the structure of variation, such as whether or not the scatter ellipsoids of several groups have major axes that are parallel. Principal components, used in many allometric studies, can "account for" or "take up" variation, but do not "cause" or "explain" it. Interpretation and explanation are extrinsic to the analyses, and they consist of arguments about biological processes producing the observed patterns of variation, e.g., growth dynamics or evolutionary constraints.

A very different framework underlies the *factor analytic approach*, which starts with an explicit model of the origin of variation in measurement variables (see Bookstein et al., 1985; Bookstein, 1989, 1991). Factors are formally included in the analysis as causes of covariation among several morphometric variables; explanation is therefore an intrinsic part of the analysis. A path model is constructed according to biological knowledge, and specifies a hypothesis of relationships between factors and observed variables (Wright,

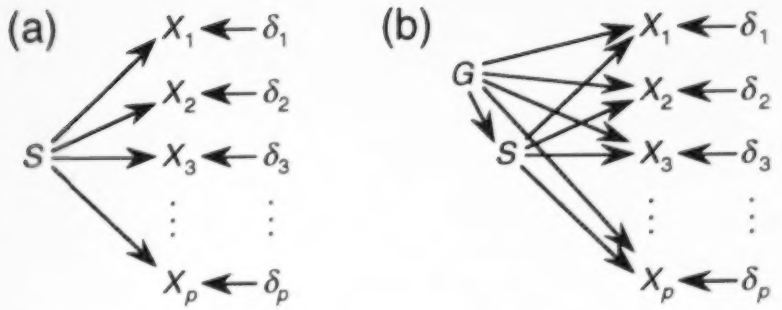


Fig. 2-8. Path models in allometry. (a) Path diagram for simple allometry in a single group. All covariation among variables is caused by a general size factor *S*. Residual variation (δ) is uncorrelated among variables. (b) Allometric variation in two groups. Both groups share the same general size factor, and therefore also the same within-group covariance structure. Group differences are determined by the group factor *G* in two different ways: directly as group shape differences (arrows from *G* to the variables), or as group size difference via the general size factor (arrow from *G* to *S*).

1968; Bookstein et al., 1985; Loehlin, 1987; Zelditch, 1987; Marcus, 1990; Bookstein, 1991). In a path model of allometry, *general size* is a factor, or latent variable, simultaneously affecting all morphometric measurements and causing the covariance among them (Fig. 2-8a). As a latent variable, size cannot be measured directly, but it can be estimated. This is usually done using the within-group PC1 (Bookstein et al., 1985; Rohlf and Bookstein, 1987).

Hopkins (1966) proposed a similar factor model of allometry. The observed covariance matrix **S** of log-transformed characters is composed of two parts, **S** = **T** + **D**. The matrix **T**, which reflects systematic covariance, is of rank one. Therefore, it has only one principal component, which corresponds to the allometric axis, or to the general size factor in figure 2-8a. The PC coefficients of **T** (which cannot be observed) are proportional to the factor loadings of general size. The matrix **D** stands for the residual variation, which is assumed to be uncorrelated among variables; therefore, **D** is a diagonal matrix. The structure of **D** is crucial for the choice of methods to estimate the parameters of the model. The PC1 of **S** is only appropriate to estimate the size factor if the diagonal elements of **D** are equal (Hopkins, 1966); otherwise, the size factor should be estimated from the off-diagonal elements (see Bookstein, 1991). If all variables are highly correlated to each other, as in many morphometric data sets, the PC1 is a reasonable estimator.

If there is more than one group of specimens, one or more additional factors for group differences (*G* in Fig. 2-8b) explain differences between two or more groups (Bookstein et al., 1985; Rohlf and Bookstein, 1987; Bookstein, 1991). Group factors can affect the measurements through general size or in a size-invariant manner. Group size differences (arrow from *G* to *S* in Fig. 2-8b) cause shifts along the growth axis, and correspond to ontogenetic scaling, whereas size-invariant differences (arrows from *G* to the variables in Fig. 2-8b) corre-

spond to lateral shifts of trajectories (Shea, 1985, 1992; Klingenberg and Spence, 1993). This is the path model for the *shearing* procedure, which was originally introduced by Humphries et al. (1981), and later reformulated by Bookstein et al. (1985), Rohlf and Bookstein (1987), and Bookstein (1991). The main purpose of this procedure is to obtain factor loadings interpretable as path coefficients (Rohlf and Bookstein, 1987), rather than ordination. The geometric basis for the shear is more complex than for Burnaby's procedure because it is not just a rigid rotation (see Humphries et al., 1981; Bookstein et al., 1985; Rohlf and Bookstein, 1987). As a consequence, it does not conserve the spatial relationships among data points and is difficult to use, e.g., to quantify lateral transposition of growth trajectories. Applications of the shear include Bookstein et al. (1985), Strauss (1985), Voss et al. (1990), and Voss and Marcus (1992).

Several studies have used factor analysis to investigate more complex models of correlation or covariance structure among morphometric variables (Bookstein et al., 1985; Zelditch, 1987, 1988; Zelditch and Carmichael, 1989; Marcus, 1990). In addition to general size, these models include factors explaining joint variation in groups of variables that are developmentally or functionally related. An alternative procedure (Cowley and Atchley, 1990; Paulsen and Nijhout, 1993) uses hypotheses about relations among characters to predict the pattern of a correlation matrix, and then compares these to observed correlation matrices using randomization tests (Cheverud et al., 1989; Manly, 1991). These models, however, are beyond the scope of allometry.

Yet another, more general method of characterizing size was introduced by Mosimann (1970), when he defined *standard size variables*. Any positive, real-valued function $G(\mathbf{x})$ of a vector of measurements $\mathbf{x}$ is a standard size variable, if multiplication of each measurement by a constant a results in an a-fold value of the size function, i.e., $G(a\mathbf{x}) = aG(\mathbf{x})$. This condition ensures that the variable scales as a linear dimension. A standard size variable transformed to logarithms is called a log-size variable. A class of log-size variables important for multivariate allometry is defined as linear combinations of log-transformed measurements, i.e., $\log G(\mathbf{x}) = \sum b_i(\log x_i)$, with $\sum b_i = 1$ (Mosimann and James, 1979; Darroch and Mosimann, 1985). Rescaling the PC coefficients for ontogenetic allometry so that they sum up to unity yields a log-size variable that indicates each specimen's position along the growth trajectory (Klingenberg and Zimmermann, 1992*b*; Klingenberg and Spence, 1993). Similar measures of size, but without rescaling, were used by Creighton and Strauss (1986), Strauss (1990*b*) and Voss and Marcus (1992), among others.

If size alone is of interest, the choice of a size measure often does not matter very much. In many morphometric data sets, the variables and the size measures derived from them are highly correlated among one another. Hence, different size measures may produce different scaling factors, due to allometry, but they will yield basically the same ordering from small to large specimens, and similar size differences between them. In studies of allometry based on either PCA or factor analysis, "shape" often does not appear explicitly at all, or if it does, it is used in a sense very different from everyday language (e.g., Bookstein, 1989). In those cases, "shape" is usually (though rarely explicitly) defined as "everything that is not size." It is through this notion of "shape" that the choice of a size measure matters for morphometric studies: because "size"

often takes up a large fraction of the total variation in a data set, relatively small changes in the size measure produce proportionally large changes in what remains after "size" is removed.

In multivariate allometry based on PCA, the second and subsequent PCs often are interpreted as "shape scores." Nevertheless, they do not reflect a geometric concept of shape. If two specimens, differing in size, have the same "shape scores," they can be interpreted as geometrically similar only if the corresponding size vector is isometric; otherwise, there are allometric changes in shape (see also Bookstein, 1989). To separate size and geometric shape, Somers (1986, 1989) proposed a size-constrained version of PCA in which variation in the direction of an isometric vector is removed first. Unfortunately, Somers used the correlation instead of the covariance matrix, thereby removing not only isometric size from the data, but also all allometric variation (see above). As an alternative procedure to achieve Somers's original objective, Burnaby's procedure can be used to eliminate a vector representing isometric variation, which is mathematically equivalent to performing a PCA on the covariance matrix of doubly centered data (Somers, 1989) or to the "principal components of shape" proposed by Darroch and Mosimann (1985; see also Jungers et al., 1988). Notice, however, that removing isometric size adjusts only for variability in overall size itself, but not for any size- or age-related shape variation. For example, although Barbie dolls are smaller than many other dolls clearly representing infants, it is easily recognizable that they are modeled after human adults.

Mosimann (1970) presented a definition of a shape space, based on geometric similarity. Each vector of measurements $\mathbf{x}$, divided by a standard size variable $G(\mathbf{x})$, constitutes a measure of shape. Darroch and Mosimann (1985) developed principal components and canonical variates for the space of these shape measures, and applied them to two examples. Further applications are found in Mosimann and James (1979) and Jungers et al. (1988). In this framework, allometry exists if variations in size and shape are associated; isometry means that variation in size and shape are statistically independent. Mosimann's theory of size and shape links some aspects of multivariate allometry to the landmark-based methods of geometric morphometrics (Bookstein, 1989, 1991, 1993). The underlying concept of allometry, however, differs fundamentally from the other frameworks presented here, as it abandons the straight-line relationship among log-transformed variables, which is the basis of allometry as devised by Huxley (1932). In this point, Mosimann's concept is closer to the much broader notion of allometry adopted by Gould (1966), who characterized it as "the study of size and its consequences."

Whereas Mosimann's (1970) approach, although based on considerations of geometric similarity, still uses vectors of length measurements, geometric morphometrics goes one step further and analyzes shape as *geometric configurations* of morphological landmarks (e.g., Rohlf, 1990; Bookstein, 1991, 1993; Rohlf and Marcus, 1993; other chapters in this volume). The strong emphasis on shape in geometric morphometrics is reflected in two recent definitions of morphometrics, characterizing it as "the quantitative description, analysis, and interpretation of shape and shape variation in biology" (Rohlf 1990) and as "the geometrically reified description of effects on geometric shape" (Bookstein, 1993) without even mentioning size. Clearly, geometric morphometrics presents a dramatically different framework for allometry. Variation in

size is removed from the data (by the two-point registration for Bookstein's shape coordinates, or by standardizing for centroid size; Bookstein, 1991; Rohlf, 1993) and shape changes alone are included in the analysis. Allometry can be assessed by combining the results from shape analysis with additional information, either directly by nonlinear regression of "shape scores" (relative warps, shape coordinates, or procrustes residuals) on a measure of size (Bookstein, 1991; Walker, 1993), or by subdividing specimens into size classes (MacLeod and Kitchell, 1990) or age groups (Reilly, 1990; Bookstein, 1991; Zelditch et al., 1992) and comparing their mean shapes. Allometric variation over an extended size range, as it occurs in many growth studies, often leads to highly nonlinear trajectories (Bookstein, 1991; Zelditch et al., 1992; Walker, 1993).

The choice of methods for a particular study depends on what questions a study is supposed to answer. The results obtained from analyses using the geometric methods can be interpreted directly in terms of shape. Here, "shape" is used in its intuitive sense, meaning a geometric configuration. The disadvantage of these methods, however, is the complexity of allometric relations. For example, a procedure analogous to Burnaby's technique to adjust for shape differences due to allometric growth, would have to use nonlinear regression of shape measures on overall size. On the other hand, the results of analyses using distance data are more difficult to describe in everyday language; graphical displays like figure 2-7 are abstractions rather than pictures of real organisms. If the notion of "shape" is used at all, it denotes the relative sizes of parts of the organism. The advantage of methods using log-transformed distances is that they often fit linear models due to their relationship to growth dynamics, which was used by Huxley (1932) to justify his formula for simple allometry.

From an extreme point of view, the configuration of morphological landmarks of an organism could be considered as merely an epiphenomenon of the growth processes affecting the tissues between the landmarks. Ideally, therefore, morphometric methods should be based on models of biological processes rather than geometrical or statistical considerations (e.g., Sattler, 1992). While this view is correct in principle, our knowledge of the mechanisms involved in developmental processes is incomplete even for simple and well-studied experimental systems (e.g., Atchley and Hall, 1991). For less well-known organisms and for more complex problems, such as evolutionary comparisons, landmark configurations or length measurements must be used as the basis for our understanding of organismic form.

References

Airoldi, J.-P., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of Microtus californicus and M. ochrogaster (Mammalia, Rodentia). Journal of Zoology, London 216: 21–36.

Anderson, T. W. 1963. Asymptotic theory for principal component analysis. *Annals of Mathematical Statistics* **34:** 122-148.

Anstey, R. L. 1987. Astogeny and phylogeny: evolutionary heterochrony in Paleozoic bryozoans. *Paleobiology* **13:** 20–43.

Arnold, S. J. 1992. Constraints on phenotypic evolution. American Naturalist 140:S85–S107.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. *Biological Reviews* **66:** 101–157.

Bernays, E. A. 1986. Diet-induced head allometry among foliage-chewing insects and its importance for graminivores. *Science* **231:** 495–497.

Blackstone, N. W. 1987. Specific growth rates of parts in a hermit crab *Pagurus longicarpus*: A reductionist approach to the study of allometry. *Journal of Zoology, London* **211:** 531–545.

Boag, P. T. 1984. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galapagos. *Journal of Zoology, London* **204:** 413–441.

Boitard, M., J. Lefebvre, and M. Solignac. 1982. Analyse en composantes principales de la variabilité de taille, de croissance et de conformation des espèces du complexe *Jaera albifrons* (Crustacés Isopodes). *Cahiers de Biologie Marine* **23:** 115–142.

Bookstein, F. L. 1989. "Size and shape": A comment on semantics. *Systematic Zoology* **38:** 173–180.

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge, U.K.

Bookstein, F. L. 1993. A brief history of the morphometric synthesis. Pages 15–40. In *Contributions to Morphometrics* (L. F. Marcus, E. Bello, and A. García-Valdecasas, eds.). Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Bookstein, F. L., B. Chernoff, R. L. Elder, J. M. Humphries, Jr., G. R. Smith, and R. E. Strauss. 1985. *Morphometrics in Evolutionary Biology: The Geometry of Size and Shape Change, with Examples from Fishes*. Academy of Natural Sciences of Philadelphia, Special Publication 15.

Bryant, P. 1984. Geometry, statistics, probability: Variations on a common theme. *American Statistician* **38:** 38–48.

Burnaby, T. P. 1966. Growth-invariant discriminant functions and generalized distances. *Biometrics* **22:** 96–110.

Buss, L. W., and N. W. Blackstone. 1991. An experimental exploration of Waddington's epigenetic landscape. *Philosophical Transactions of the Royal Society of London B* **332:** 49–58.

Cane, W. P. 1993. The ontogeny of postcranial integration in the common tern, *Sterna hirundo*. Evolution 47:1138–1151.

Cheverud, J. M. 1982. Relationships among ontogenetic, static, and evolutionary allometry. *American Journal of Physical Anthropology* **59:** 139–149.

Cheverud, J. M., G. P. Wagner, and M. M. Dow. 1989. Methods for the comparative analysis of variation patterns. *Systematic Zoology* **38:** 201–213.

Cock, A. G. 1966. Genetical aspects of metrical growth and form in animals. Quarterly Review of Biology 41: 131–190.

Corner, B. D., and J. T. Richtsmeier. 1991. Morphometric analysis of craniofacial growth in *Cebus apella*. *American Journal of Physical Anthropology* **84:** 323–342.

Corti, M., R. S. Thorpe, L. Sola, V. Sbordoni, and S. Cataudella. 1988. Multivariate morphometrics in aquaculture: A case study of six stocks of the common carp (*Cyprinus carpio*) from Italy. *Canadian Journal of Fisheries and Aquatic Sciences* **45:** 1548–1554.

Cowley, D. E., and W. R. Atchley. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. *American Naturalist* **135:** 242–268.

Cowley, D. E., and W. R. Atchley. 1992. Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. *Evolution* **46:** 495–518.

Creighton, G. K., and R. E. Strauss. 1986. Comparative patterns of growth and development in cricetine rodents and the evolution of ontogeny. *Evolution* **40:** 94–106.

Cuzin-Roudy, J. 1975. Étude de la variabilité et de l'allométrie de taille chez *Notonecta maculata* Fabricius (Insectes, hétéroptères), par les méthodes classiques et par la méthode des composantes principales. *Archives de Zoologie Expérimentale et Générale* **116:** 173–227.

Cuzin-Roudy, J., and Ph. Laval. 1975. A canonical discriminant analysis of post-embryonic development in *Notonecta maculata* Fabricius (Insecta: Heteroptera). *Growth* **39:** 251–280.

Darroch, J. N., and J. E. Mosimann. 1985. Canonical and principal components of shape. *Biometrika* **72:** 241–252.

Daudin, J. J., C. Duby, and P. Trecourt. 1988. Stability of principal component analysis studied by the bootstrap method. *Statistics* **19:** 241–258.

Davies, R. G., and V. Brown. 1972. A multivariate analysis of postembryonic growth in two species of *Ectobius* (Dictyoptera: Blattidae). *Journal of Zoology, London* **168:** 51–79.

Diaconis, P., and B. Efron. 1983. Computer-intensive methods in statistics. *Scientific American* **248** (May): 116–130.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. *American Statistician* **37:** 36–48.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* **1:** 54–77.

Efron, B., and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* **125:** 1–15.

Flury, B. 1988. *Common Principal Components and Related Multivariate Models*. Wiley, New York.

Flury, B., and H. Riedwyl. 1988 *Multivariate Statistics: A Practical Approach*. Chapman and Hall, London.

Garland, T., Jr., P. H. Harvey, and A. R. Ives. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Systematic Biology* **41:** 18–32.

Gibson, A. R., A. J. Baker, and A. Moeed. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): An application of the jackknife to principal component analysis. *Systematic Zoology* **33:** 408–421.

Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. *Biological Reviews* **41:** 587–640.

Gould, S. J. 1975. Allometry in primates, with emphasis on scaling and the evolution of the brain. *Contributions to Primatology* **5:** 244–292.

Gould, S. J. 1989. A developmental constraint in *Cerion*, with comments on the definition and interpretation of constraint in evolution. *Evolution* **43:** 516–539.

Gower, J. C. 1976. Growth-free canonical variates and generalized inverses. *Bulletin of the Geological Institutions of the University of Uppsala*, N. S. **7:** 1–10.

Günther, B. 1975. Dimensional analysis and theory of biological similarity. *Physiological Reviews* **55:** 659–699.

Hills, M. 1982. Bivariate versus multivariate allometry: A note on a paper by Jungers and German. *American Journal of Physical Anthropology* **59:** 321–322.

Hopkins, J. W. 1966. Some considerations in multivariate allometry. *Biometrics* **22:** 747–760.

Humphries, J. M., F. L. Bookstein, B. Chernoff, G. R. Smith, R. L. Elder, and S. G. Poss. 1981. Multivariate discrimination by shape in relation to size. *Systematic Zoology* **30:** 291–308.

Huxley, J. S. 1932. *Problems of Relative Growth*. Methuen, London. Reprinted, Dover Publications, New York, 1972.

Huxley, J. S., and G. Teissier. 1936. Terminology of relative growth. *Nature* **137:** 780–781.

Jackson, J. E. 1990. *A User's Guide to Principal Components*. Wiley, New York.

Jeune, B., and R. Sattler. 1992. Multivariate analysis in process morphology of plants. *Journal of theoretical Biology* **156:** 147–167.

Jobson, J. D. 1992. *Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate Methods*. Springer-Verlag, New York.

Johnson, R. A., and D. W. Wichern. 1988. *Applied Multivariate Statistical Analysis*. 2nd edition. Prentice-Hall, Englewood Cliffs, NJ.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. *Biometrics* **19:** 497–499.

Jolliffe, I. T. 1986. *Principal Component Analysis*. Springer-Verlag, New York.

Jones, C. S. 1992. Comparative ontogeny of a wild cucurbit and its derived cultivar. *Evolution* **46:** 1827–1847.

Jones, C. S. 1993. Heterochrony and heteroblastic leaf development in two subspecies of *Cucurbita argyrosperma* (Cucurbitaceae). *American Journal of Botany* **80:** 778–795.

Jungers, W. L., and R. Z. German. 1981. Ontogenetic and interspecific skeletal allometry in nonhuman primates: Bivariate versus multivariate analysis. *American Journal of Physical Anthropology* **55:** 195–202.

Jungers, W. L., T. M. Cole, III, and D. W. Owsley. 1988. Multivariate analysis of relative growth in the limb bones of Arikara Indians. *Growth, Development and Aging* **52:** 103–107.

Kampny, C. M., T. A. Dickinson, and N. G. Dengler. 1993. Quantitative comparison of floral development in *Veronica chamaedrys* and *Veronicastrum virginicum* (Scrophulariaceae). *American Journal of Botany* 80: 449–460.

Katz, M. J. 1980. Allometry formula: A cellular model. *Growth* **44:** 89–96.

Klingenberg, C. P. 1992. Multivariate morphometrics of geographic variation of *Gerris costae* (Heteroptera: Gerridae) in Europe. *Revue suisse de Zoologie* **99:** 11–30.

Klingenberg, C. P., and R. Froese. 1991. A multivariate comparison of allometric growth patterns. *Systematic Zoology* **40:** 410–419,

Klingenberg, C. P., and J. R. Spence. 1993. Heterochrony and allometry: Lessons from the waterstrider genus *Limnoporus*. *Evolution*. 47:1834-1853.

Klingenberg, C. P., and M. Zimmermann. 1992*a*. Static, ontogenetic, and evolutionary allometry: A multivariate comparison in nine species of water striders. *American Naturalist* **140:** 601–620.

Klingenberg, C. P., and M. Zimmermann. 1992*b*. Dyar's rule and multivariate allometric growth in nine species of waterstriders (Heteroptera: Gerridae). *Journal of Zoology, London* **227:** 453–464.

LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. *Annual Review of Ecology and Systematics* **20:** 97–117.

Laird, A. K. 1965. Dynamics of relative growth. *Growth* **29:** 249–263.

Laird, A. K., A. D. Barton, and S. A. Tyler. 1968. Growth and time: An interpretation of allometry. *Growth* **32:** 347–354.

Leamy, L. and W. Atchley. 1984*a*. Static and evolutionary allometry of osteometric traits in selected lines of rats. *Evolution* **38:** 47–54.

Leamy, L., and W. R. Atchley. 1984*b*. Morphometric integration in the rat (*Rattus* sp.) scapula. *Journal of Zoology, London* **202:** 43–56.

Leamy, L., and D. Bradley. 1982. Static and growth allometry of morphometric traits in randombred house mice. *Evolution* **36:** 1200–1212.

Lessa, E. P., and J. L. Patton. 1989. Structural constraints, recurrent shapes, and allometry in pocket gophers (genus *Thomomys*). *Biological Journal of the Linnean Society*. **36:** 349–363.

Livezey, B. C. 1989. Morphometric patterns in Recent and fossil penguins (Aves, Sphenisciformes). *Journal of Zoology, London* **219:** 269–307.

Loehlin, J. C. 1987. *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*. Lawrence Erlbaum Associates, Hillsdale, NJ.

MacLeod, N., and J. A. Kitchell. 1990. Morphometrics and evolutionary inference: A case study involving ontogenetic and developmental aspects of foraminiferal evolution. Pages 283–299. In *Proceedings of the Michigan Morphometrics Workshop* (F. J. Rohlf and F. L.

Bookstein, eds.). Special Publication No. 2. The University of Michigan Museum of Zoology, Ann Arbor, Michigan.

Maddison, W. P., and D. R. Maddison. 1992. *MacClade: Analysis of Phylogeny and Character Evolution*. Sinauer Associates, Sunderland, Mass.

Manly, B. F. J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London.

Marcus, L. F. 1990. Traditional morphometrics. Pages 77–122. In *Proceedings of the Michigan Morphometrics Workshop* (F. J. Rohlf and F. L. Bookstein, eds.). University of Michigan Museum of Zoology, Special Publication 2.

Marcus, L. F. 1993. Some aspects of multivariate statistics for morphometrics. Pages 95–130. In Contributions to Morphometrics (L. F. Marcus, E. Bello, and A. García-Valdecasas, eds.). Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Martins, E. P., and T. Garland, Jr. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: A simulation study. *Evolution* **45:** 534–557.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, and L. Wolpert. 1985. Developmental constraints and evolution. *Quarterly Review of Biology* **60:** 265–287.

McGillivray, W. B. 1985. Size, sexual size dimorphism, and their measurement in Great Horned Owls in Alberta. *Canadian Journal of Zoology* **63:** 2364–2372.

McLellan, T. 1993. The roles of heterochrony and heteroblasty in the diversification of leaf shapes in *Begonia dregei* (Begoniaceae). *American Journal of Botany* **80:** 796–804.

Meyer, A. 1990. Morphometrics and allometry in the trophically polymorphic cichlid fish, *Cichlastoma citrinellum*: Alternative adaptations and ontogenetic changes in shape. *Journal of Zoology, London* **221:** 237–260.

Mosimann, J. E. 1970. Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions. *Journal of the American Statistical Association* **65:** 930–945.

Mosimann, J. E., and F. C. James. 1979. New statistical methods for allometry with application to Florida red-winged blackbirds. *Evolution* **33:** 444–459.

Owen, J. G., and M. A. Chmielewski. 1985. On canonical variates analysis and the construction of confidence ellipses in systematic studies. *Systematic Zoology* **34:** 366–374.

Pagel, M. D., and P. H. Harvey. 1988. Recent developments in the analysis of comparative data. *Quarterly Review of Biology* **63:** 413–440.

Pandolfi, J. M. 1988. Heterochrony in colonial marine animals. Pages 135–158. In *Heterochrony in Evolution: A Multidisciplinary Approach* (M. L. McKinney, ed.). Plenum Press, New York.

Patton, J. L., and P. V. Brylski. 1987. Pocket gophers in alfalfa fields: causes and consequences of habitat-related body size variation. *American Naturalist* **130:** 493–506.

Paulsen, S. M., and H. F. Nijhout. 1993. Phenotypic correlation structure among elements of the color pattern in *Precis coenia* (Lepidoptera: Nymphalidae). *Evolution* **47:** 593–618.

Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine and Journal of Science*, 6th series **2:** 559–572.

Pimentel, R. A. 1979. *Morphometrics: The Multivariate Analysis of Biological Data*. Kendall/Hunt, Dubuque, Iowa.

Reeve, E. C. R., and J. S. Huxley. 1945. Some problems in the study of allometric growth. Pages 121–156. In *Essays on Growth and Form presented to D'Arcy Wentworth Thompson*. (W. E. Le Gros Clark and P. B. Medawar, eds.). Clarendon Press, Oxford.

Reilly, S. M. 1990. Comparative ontogeny of cranial shape in salamanders using resistant fit theta rho analysis. Pages 311–321. In *Proceedings of the Michigan Morphometrics Workshop* (F. J. Rohlf and F. L. Bookstein, eds.). Special Publication No. 2. The University of Michigan Museum of Zoology, Ann Arbor, Michigan.

Reiss, M. J. 1989. *The Allometry of Growth and Reproduction*. Cambridge University Press, Cambridge, U.K.

Reyment, R. A. 1991. *Multidimensional Palaeobiology*. Pergamon Press, Oxford.

Reyment, R. A., and C. F. Banfield. 1976. Growth-free canonical variates applied to fossil foraminifers. *Bulletin of the Geological Institutions of the University of Uppsala*, N. S. **7:** 11–21.

Reyment, R. A., R. E. Blackith, and N. A. Campbell. 1984. *Multivariate Morphometrics*. 2nd edition. Academic Press, London.

Riska, B. 1981. Morphological variation in the horseshoe crab *Limulus polyphemus*. *Evolution* **35:** 647–658.

Riska, B. 1986. Some models for development, growth, and morphometric correlation. *Evolution* **40:** 1303–1311.

Riska, B. 1989. Composite traits, selection response, and evolution. *Evolution* **43:** 1172–1191.

Rohlf, F. J. 1990. Morphometrics. *Annual Review of Ecology and Systematics* **21:** 299–316.

Rohlf, F. J. 1993. Relative warp analysis and an example of its application to mosquito wings. Pages 131–159. In *Contributions to Morphometrics* (L. F. Marcus, E. Bello, and A. García-Valdecasas, eds.). Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Rohlf, F. J., and F. L. Bookstein. 1987. A comment on shearing as a method for "size correction." *Systematic Zoology* **36:** 356–367.

Rohlf, F. J., and L. F. Marcus. 1993. A revolution in morphometrics. *Trends in Ecology and Evolution* **8:** 129–132.

Sattler, R. 1992. Process morphology: Structural dynamics in development and evolution. *Canadian Journal of Botany* **70:** 708–714.

Shea, B. T. 1985. Bivariate and multivariate growth allometry: statistical and biological considerations. *Journal of Zoology, London* **206:** 367–390.

Shea, B. T. 1992. Developmenal perspective on size change and allometry in evolution. *Evolutionary Anthropology* **1:** 125–134.

Shea, B. T., R. E. Hammer, R. L. Brinster, and M. R. Ravosa. 1990. Relative growth of the skull and postcranium in giant transgenic mice. *Genetical Research, Cambridge* **56:** 21–34.

Sinervo, B. 1993. The effect of offspring size on physiology and life history. *BioScience* **43:** 210–218.

Smith, L. D., and A. R. Palmer. 1994 Effects of manipulated diet on size and performance of brachyuran crab claws. Science 264:710–712.

Smith, R. J. 1993. Logarithmic transformation bias in allometry. *American Journal of Physical Anthropology* **90:** 215–228.

Solignac, M., M.-L. Cariou, and M. Wimitzki. 1990. Variability, specificity and evolution of growth gradients in the species complex *Jaera albifrons* (Isopoda, Asellota). *Crustaceana* **59:** 121–145.

Somers, K. M. 1986. Multivariate allometry and removal of size with principal components analysis. *Systematic Zoology* **35:** 359–368.

Somers, K. M. 1989. Allometry, isometry and shape in principal component analysis. Systematic Zoology 38: 169–173.

Sprent, P. 1972. The mathematics of size and shape. *Biometrics* **28:** 23–37.

Stauffer, D. F., E. O. Garton, and R. K. Steinhorst. 1985. A comparison of principal components from real and random data. *Ecology* **66:** 1693–1698.

Strauss, R. E. 1985. Evolutionary allometry and variation in body form in the South American catfish genus *Corydoras* (Callichthyidae). *Systematic Zoology* **34:** 381–396.

Strauss, R. E. 1990*a*. Patterns of quantitative variation in lepidopteran wing morphology: The convergent groups Heliconiinae and Ithomiinae (Papilionoidea: Nymphalidae). *Evolution* **44:** 86–103.

Strauss, R. E. 1990*b*. Heterochronic variation in the developmental timing of cranial ossifications in poecilid fishes (Cyprinodontiformes). *Evolution* **44:** 1558–1567.

Strauss, R. E., and F. L. Bookstein. 1982. The truss: Body form reconstructions in morphometrics. *Systematic Zoology* **31:** 113–135.

Teissier, G. 1955. Allométrie de taille et variabilité chez *Mala squinado*. *Archives de Zoologie Expérimentale et Générale* **92:** 221–264.

Thomas, R. D. K., and W.-E. Reif. 1993. The skeleton space: A finite set of organic designs. *Evolution* **47:** 341-360.

Thorpe, R. S. 1983. A review of the numerical methods for recognising and analysing racial differentiation. Pages 404–423. In *Numerical Taxonomy* (J. Felsenstein, ed.). Springer-Verlag, Berlin.

Thorpe, R. S., and M. Baez. 1987. Geographic variation within an island: Univariate and multivariate contouring of scalation, size, and shape of the lizard *Gallotia galloti*. *Evolution* **41:** 256–268.

Voss, R. S., L. F. Marcus, and P. Escalante. 1990. Morphological evolution in muroid rodents I. Conservative patterns of craniometric covariance and their ontogenetic basis in the neotropical genus *Zygodontomys*. *Evolution* **44:** 1568–1587.

Voss, R. S., and L. F. Marcus. 1992. Morphological evolution in muroid rodents II. Craniometric factor divergence in seven neotropical genera, with experimental results from *Zygodontomys*. *Evolution* **46:** 1918–1934.

Walker, J. A. 1993. Ontogenetic allometry of threespine stickleback body form using landmark-based morphometrics. Pages 193–214. In *Contributions to Morphometrics* (L. F. Marcus, E. Bello, and A. García-Valdecasas, eds.). Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Wheeler, D. E. 1991. The developmental basis of worker caste polymorphism in ants. American Naturalist 138: 1218–1238.

Wiig, Ø. 1985. Multivariate variation in feral American mink (*Mustela vison*) from Southern Norway. *Journal of Zoology, London* **206:** 441–452.

Wright, S. 1968. *Evolution and the Genetics of Populations. Vol. 1. Genetic and Biometric Foundations*. University of Chicago Press, Chicago.

Zelditch, M. L. 1987. Evaluating models of developmental integration in the laboratory rat using confirmatory factor analysis. *Systematic Zoology* **36:** 368–380.

Zelditch, M. L. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. *Evolution* **42:** 28–41.

Zelditch, M. L., F. L. Bookstein, and B. L. Lundrigan. 1992. Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer*. *Evolution* **46:** 1164–1180.

Zelditch, M. L., and A. C. Carmichael. 1989. Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon fulviventer*. *Evolution* **43:** 814–824.

APPENDIX

Morphometric data (raw values, in millimetres) for the example in the text. Four measurements were made on male *Gerris costae* from three different locations in Europe (Klingenberg, 1992). The measurements are the lengths of the thorax (Tho), the first antennal segment (Ant), and the femora of the middle and hind legs (MF and HF), corresponding to nos. 35, 43, 49, and 53 of Klingenberg (1992).

Alps				Greece				Pyrenees			
Tho	Ant	MF	HF	Tho	Ant	MF	HF	Tho	Ant	MF	HF
5.11	1.51	6.07	6.13	5.74	1.72	6.65	7.05	5.15	1.45	6.05	6.36
5.10	1.53	6.15	6.25	5.40	1.59	6.39	6.76	5.05	1.46	5.86	6.11
5.08	1.60	6.28	6.41	5.34	1.67	6.47	6.79	4.97	1.52	6.12	6.35
5.16	1.55	6.18	6.3	5.51	1.72	6.65	7.05	5.03	1.43	6.01	6.21
4.78	1.54	5.79	5.86	5.44	1.66	6.46	6.76	5.34	1.53	6.13	6.40
4.64	1.41	5.55	5.46	5.49	1.69	6.27	6.48	4.99	1.41	5.81	5.89
5.27	1.60	6.20	6.29	5.64	1.75	6.75	6.95	5.31	1.59	6.39	6.67
4.71	1.43	5.74	5.67	5.45	1.71	6.70	7.12	5.29	1.54	6.23	6.62
5.42	1.55	6.28	6.35	5.36	1.71	6.84	7.12	5.33	1.59	6.35	6.60
4.91	1.48	5.70	5.71	5.44	1.72	6.31	6.71	5.18	1.46	6.25	6.54
5.03	1.48	6.07	5.82	5.51	1.69	6.34	6.55	5.22	1.52	6.08	6.49
5.11	1.48	6.29	6.38	5.71	1.69	6.44	6.88	5.23	1.47	6.02	6.46
5.01	1.56	6.10	6.23	5.42	1.65	6.44	6.59	5.16	1.51	6.06	6.20
5.25	1.56	6.06	6.13	5.35	1.67	6.36	6.61	5.07	1.42	5.89	6.27
5.15	1.57	6.03	6.21	5.45	1.57	6.18	6.62	5.13	1.54	5.91	6.25
5.18	1.57	6.20	6.3	5.28	1.56	6.20	6.49	5.14	1.55	6.06	6.35
4.96	1.53	5.92	5.93	5.43	1.66	6.28	6.58	5.26	1.47	6.18	6.49
5.02	1.48	5.87	6.17	5.44	1.63	6.28	6.52	4.96	1.40	5.81	6.15
4.82	1.51	6.02	6.08	5.48	1.75	6.44	6.61	5.10	1.42	5.97	6.07
5.22	1.62	6.13	6.21	5.51	1.71	6.17	6.52	5.17	1.54	6.11	6.35
5.16	1.63	5.98	6.07	5.32	1.67	6.52	7.02	5.06	1.48	6.02	6.35
5.34	1.58	6.19	6.41	5.58	1.62	6.51	6.80	5.08	1.53	6.22	6.38
5.18	1.55	6.04	6.14	5.46	1.61	6.45	6.66	5.08	1.39	5.92	6.15
4.94	1.59	6.36	6.46	5.44	1.66	6.59	6.92	5.25	1.43	6.08	6.39
5.05	1.48	5.83	5.92	5.63	1.61	6.46	6.80	5.06	1.46	6.18	6.57
4.87	1.48	5.79	5.93	5.61	1.70	6.75	7.16	5.21	1.61	6.05	6.34
5.10	1.55	6.28	6.44	5.42	1.74	6.30	6.65	5.05	1.40	6.19	6.33
5.16	1.69	6.42	6.50	5.67	1.71	6.75	7.27	4.96	1.40	5.75	5.82
5.13	1.53	6.15	6.20	5.15	1.66	6.24	6.24				
4.82	1.50	5.85	5.79	5.36	1.75	6.52	6.86				
4.97	1.58	5.90	6.06	5.57	1.65	6.29	6.48				
5.02	1.56	6.17	6.25	5.62	1.74	6.57	7.06				
				5.53	1.71	6.45	6.86				

Heterochrony and Allometry: The Analysis of Evolutionary Change in Ontogeny

INTRODUCTION

Ontogeny and evolution are intimately and reciprocally interrelated. Evolutionary changes in morphological characters require changes in the developmental processes that produce the structures of interest. The study of the relationship between development and evolution has a century-old history, which has been reviewed by Gould (1977) and Hall (1992), among others. In recent years, there has been a variety of new attempts to integrate morphology, developmental and evolutionary biology, and phylogeny into a unified theory of the evolution of biological form (e.g., Liem 1991; Atkinson 1992; Raff 1992; Gilbert et al. 1996). This synthesis has resulted in the gradual emergence of the new discipline of evolutionary developmental biology (Hall 1992).

Most analyses of the relationship between ontogeny and phylogeny rely on comparisons of ontogenetic trajectories among related species. This approach is based on a metaphor that depicts development of an organism as a movement through a multidimensional space defined by its size, shape, and age. Ontogenetic trajectories are the paths along which growing organisms move through this space; they visualize developmental sequences.

Familiar representations of such trajectories are growth curves, where measurements of a trait are graphed against the age of the organisms, or bivariate allometric plots of measurements of two metric traits against each other (e.g., Huxley 1932) or of a measure of shape against size (Gould 1966). These two kinds of plots are simply projections of the trajectories from the multidimensional space onto different planes defined by age and a trait, or by a pair of traits, respectively. Therefore, growth curves and allometric plots display different aspects of the same ontogenetic sequence (Fig. 3-1).

Likewise, heterochrony and allometry deal with different aspects of the evolution of ontogenetic trajectories. Heterochrony, as it is defined by most current authors (e.g., McKinney and McNamara 1991), is concerned with evolutionary changes in rates and timing of developmental processes, and therefore explicitly incorporates time as well as morphological traits. These changes can lead to alterations of the growth trajectory in the subspace of the morphological traits, which is the realm of allometry. Allometry only refers to time implicitly, with respect to the rate at which growing organisms move through the character space (Fig. 3-1; see below for other concepts of allometry).

Heterochrony and allometry have been used extensively to study the evolution of ontogenies in a variety of organisms; comprehensive reviews on the subject have been published by Gould (1977) and McKinney and McNamara (1991). Several analytical frameworks have been proposed for heterochrony, which use mostly the same terminology, and therefore appear similar despite substantial differences in their conceptual basis and analytical procedures. Previous reviews have mostly emphasized these similarities; for instance,

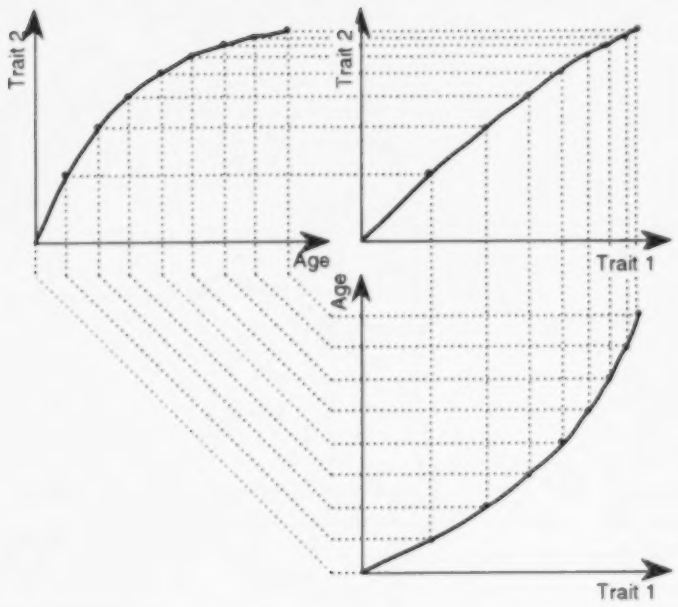


Fig. 3-1. The relationship between allometric plots and growth curves. The central metaphor of heterochrony and allometry describes the ontogeny of an organism as path through a multi-dimensional space defined by age and morphological form, exemplified here by measurements of two traits. A growth curve (top left and bottom) is a projection of this space onto the plane defined by age and a mensural trait, and explicitly characterizes the growth dynamics of that trait (here I used von Bertalanffy functions). Heterochrony pertains to evolutionary changes in these growth curves. In contrast, an allometric plot (top right) is a projection onto the plane defined by the two trait measurements. Although it results from their growth curves, the allometric plot only takes time into account indirectly, by the rate at which the organism "advances" along the ontogenetic trajectory, visualized here by the distances between dots plotted at equal time intervals. Transforming the morphometric variables (e.g., to logarithms) can often render allometric plots linear.

McKinney and McNamara (1991: 13) state that "all schemes are very similar, with concepts and terms that deviate little from the original presentation of Gould (1977)." In contrast, here I will emphasize the conceptual differences among some of the most influential contributions to this field (Gould 1977; Alberch et al. 1979; Shea 1983a, 1988, 1989; McKinney 1986, 1988; McNamara 1986; Raff and Wray 1989).

These generally underrated differences between analytical frameworks have generated confusion in terminology and methods, as they may lead to contradictory interpretations of the same evolutionary events. Miscommunication stemming from use of incompatible concepts underlies a number of current controversies, most notably the one about human heterochrony (e.g. Gould 1977; Montagu 1989; Shea 1989; McKinney and McNamara 1991; Vrba

1994). Recognizing the differences between analytical frameworks therefore is an essential first step to resolving questions about particular examples.

Resolution of conflict will be most effective if it is based on knowledge of developmental processes underlying the evolutionary changes in question. In recent years, developmental biologists have shown an increasing interest in evolution, and have applied and extended the concept of heterochrony. Heterochrony, which traditionally has been more concerned with postembryonic growth than embryonic development, is now applied throughout ontogeny from the earliest stages of embryogenesis to maturity. Moreover, advances in developmental biology are having a substantial impact on the concept of heterochrony, and have lead to significant changes from the classical formalisms of Gould (1977) or Alberch et al. (1979).

Moreover, the differences among methodological approaches also play a role in the debate about the relation between heterochrony and allometry (McKinney 1986, 1988; Blackstone and Yund 1989; McKinney and McNamara 1991; Klingenberg and Spence 1993; Godfrey and Sutherland 1995a). In this paper, I explore how heterochrony relates to allometry, and review attempts to use allometric data to infer heterochrony. Although I emphasize the differences between various approaches, I do not imply that they are mutually exclusive alternatives. To the contrary, exploration of the relationship between development and evolution will be most effective if several methods are employed in complementary ways.

Concepts of Heterochrony

The modern concept of heterochrony, which emerged after the demise of recapitulation as the prevailing dogma in the study of evolution (for a historical review, see Gould 1977), has its origin in the work of Gavin de Beer. He based the analysis of heterochrony on comparing the time when a character appears in the ontogeny of ancestors and descendants (de Beer 1958). Gould (1977) presented a detailed critique of de Beer's classification of heterochrony, and simplified his terminology. Gould also proposed a new framework for analyzing heterochrony, his "clock model", which later was revised and extended by Alberch et al. (1979) and several other authors (reviewed by McKinney and McNamara 1991). Although the clock model and the formalism of Alberch et al. (1979) are similar in their general purpose and their terminology, their application can lead to conflicting interpretations of the same evolutionary changes. Therefore, and because both formalisms are currently in use, I outline each and discuss human heterochrony as an example in which application of different formalisms can yield opposite results. In recent years, developmental biologists have started to use the concept of heterochrony, and have extended the concept to the earliest ontogenetic stages as well as to the molecular level (e.g., Raff and Wray 1989). Because the models applied in this context differ from the models of Gould and Alberch et al., I summarize this approach in a separate section.

Paedomorphosis and Peramorphosis

The explanatory power of heterochrony is based on the strong directionality inherent in ontogeny. In general, organisms acquire a more complex morpho-

logical organization as they grow larger during their ontogenies from single-celled eggs or zygotes to adult forms. During this process, virtually all properties of an organism undergo dramatic changes in a highly coordinated manner, thereby establishing a clear directionality of variation. By modifying the rates and timing of developmental processes, heterochrony translates this ontogenetic polarity into morphological variation between taxa or evolutionary lineages. In turn, inferences about these evolutionary changes can be drawn by comparing organisms at equivalent ontogenetic stages (e.g., the adult at sexual maturity) in relation to the ontogenetic directionality and a phylogenetic hypothesis.

The morphological outcomes of changes in rates and timing of development are paedomorphosis or peramorphosis; they are identified by comparisons of ancestors and descendants in relation to the ancestral ontogeny. A descendant is *paedomorphic* if its later ontogenetic stages retain characteristics from earlier stages of an ancestor. The direction of evolutionary change observed in mature stages is therefore opposite to the direction of ontogenetic change, a phenomenon called reverse recapitulation (Alberch et al. 1979). Alternatively, the descendant is *peramorphic* if its development goes beyond that of the ancestor at the standard stage, and thereby produces an exaggerated adult morphology. In this case, ontogenetic and evolutionary change have the same direction, and the descendant recapitulates the ancestral ontogeny, at least with regard to the particular characteristics under study.

Alberch et al. (1979) distinguished the morphological consequences of an ontogenetic change from its phylogenetic effect. Therefore, they introduced the term "peramorphosis" to replace the term "recapitulation" used by Gould (1977; a synonym of "gerontomorphosis", de Beer 1958). This distinction makes it possible not only to compare ancestors and descendants in terms of paedomorphosis and peramorphosis, but also contemporaneous taxa and other groups, such as individuals following different life-history tactics or the two sexes within species (see Reilly 1994; Whiteman 1994; Reilly et al. 1996). In most applications, however, the comparison is between a descendant and an ancestor inferred from fossil material or by phylogenetic analysis (e.g., Bryant and Russell 1992; Maddison and Maddison 1992).

The classical models of heterochrony are based on the concept of dissociability of maturation, growth, and development. Maturation is the progression through the various stages of ontogeny, which can be defined by morphological features or life-history changes (e.g., the pupal stage of holometabolous insects, reproductive maturity). Gould (1977:235–236) argued that growth, the increase in size, should be distinguished from development, the ontogenetic change in shape. He defined the term development to denote all shape changes, including those that are allometric consequences of growth, and stated that it should be separated from growth, which exclusively consists of the isometric component of "size increase with geometric similarity" (p. 235). Related concepts of size and shape underlie a number of approaches in morphometrics (e.g., Mosimann 1970; Bookstein 1991; Richtsmeier and Lele 1993; Jungers et al. 1995).

This notion of shape, based on geometric similarity, implies a distinction between increase in size and the allometric shape changes accompanying it. The primary justification Gould (1977) gave for this separation was the construction of his clock model (see below), in which size and shape are separate enti-

ties. As biological evidence, he cited Novák (1966), who proposed that insect larvae grow isometrically in the presence of juvenile hormone (pp 115, 157, 167, 180), and that allometric growth of adult structures (e.g., imaginal discs or wing pads) only occurs in the absence of juvenile hormone (see Nijhout 1994a). This argument is flawed because of the complete lack of quantitative data supporting isometric growth (Novák did not cite a single quantitative study in this context). Allometric growth is pervasive during the larval stage of insects (e.g. Matsuda 1961; Blackith et al. 1963; Brown and Davies 1972; Davies and Brown 1972; Klingenberg and Zimmermann 1992; Klingenberg and Spence 1993; Klingenberg 1996b), and I am not aware of any example of truly isometric growth. Furthermore, recent studies in developmental biology also have emphasized the intimate link between growth and pattern formation:

> Relationships between growth control and pattern formation [are] a general feature of epimorphic systems. A mechanistic linkage between the growth of a structure and the processing of its patterning system ... would help prevent these two aspects of morphogenesis from becoming uncoupled, i.e. prevent growth from occurring faster than patterning or vice-versa. (Duboule 1994:136.)

Therefore, although the separation of growth as isometric size increase from all shape changes agrees with our intuitive concept of size and shape based on geometric similarity, it does not reflect a corresponding separation of underlying biological processes.

In accordance with the idea of dissociation of size and shape, Gould (1977) and Alberch et al. (1979) applied the concepts of paedomorphosis and peramorphosis exclusively to measures of shape, but not to size (see also Godfrey and Sutherland 1995a, 1995b). Therefore, Gould (1977:256) and Alberch et al. (1979: Table 1) specified that proportional dwarfism and proportional giantism do not produce paedomorphosis or peramorphosis. Nevertheless, because allometric growth is virtually ubiquitous, shape and size are tightly linked, and shape change accompanies every change in size. *Proportional* dwarfism or giantism therefore are rare; I do not know any quantitative study showing either of them unambiguously. Hence, in most cases the size of an organism can provide valid information about ontogenetic polarity.

Organismal form is an intrinsically multivariate concept, whether it is characterized through a geometric concept of shape (e.g., Mosimann 1970; Bookstein 1991) or by the relative sizes of parts (e.g., Klingenberg 1996a). Morphometric variables extracted from different parts of an organism can thus show different ontogenetic trends, and their evolutionary history may be either independent or linked to that of other such parts. Therefore, the results of heterochronic changes can differ depending on the traits under consideration—size or shape, or different shape measures—and statements such as "the descendant species is paedomorphic" are meaningless unless it is clear to which traits they refer. Paedomorphosis and peramorphosis are relative terms, and therefore depend on the organisms being compared (e.g., an ancestor and a descendant) and on the measure of shape or size used as a criterion.

The classical models of heterochrony are based on the implicit assumption that there is an unambiguous ontogenetic polarity. This means that the measures of size and shape should increase or decrease monotonically in both an-

cestral and descendant ontogenies. For size, this is fulfilled for the vast majority of organisms, because they grow but usually do not shrink. For many shape measures, however, there may be a reversal in the direction of ontogenetic change. Then, it is not always possible to interpret evolutionary changes as paedomorphosis or peramorphosis, because the basis for the comparison changes depends on the part of the ancestral ontogeny that serves as the standard for comparison (see Dommergues 1986; Dommergues and Meister 1989). This situation may be quite common, especially when biphasic growth is involved. Imagine an example in which a trait y first grows with positive allometry and later with negative allometry relative to another variable *x*. Then the shape measure defined by their ratio, *y*/*x*, will first increase and then decrease with age. Depending on whether one chooses the first or the second growth phase as the standard for comparison, the same outcome can be interpreted as either paedomorphosis or peramorphosis. This is at least a partial explanation for unusual and apparently paradoxical heterochronies in conjunction with biphasic growth (Gould 1977, footnote on p 365; Shea 1989:82; Vrba 1994). Also, it is clear that paedomorphosis and peramorphosis are inapplicable for shape if the ancestor shows isometric growth, that is, increase in size without concomitant shape change. The importance of analyzing complete growth trajectories cannot be overemphasized in this context (Dommergues 1986). In most cases, a different choice of the shape variable will avoid this difficulty.

This discussion of paedomorphosis and peramorphosis shows that these terms strongly depend on the context in which they are used: different choices of the morphological feature and of the ontogenetic stages included may produce contradictory results. The methods and underlying theories of measuring shape and size of organisms are critical determinants of the results of each analysis. It is therefore important to state this context explicitly in any empirical study.

The Clock Model

The ontogeny of morphological form can be described as a sequence of coordinated changes associated with age that affect size and shape of organisms. Evolutionary modifications in ontogeny can affect the size, shape, and the age at which the organism attains any particular developmental stage. Ancestral relationships among size, shape, and age can either be conserved or modified; the latter possibility, dissociation, is the focus of Gould's (1977) approach to heterochrony.

Gould (1977) proposed his clock model as a graphical device to display and compare the ontogenies of ancestors and descendants. The clock has three scales: one each for a measure of size, a measure of shape, and age (Fig. 3-2A). The scales are calibrated for the ancestor, and the hands of the clock display the ontogeny of the descendant, thereby revealing possible differences in development. This model also serves as the basis for a classification system of heterochronic changes (Fig. 3-2B).

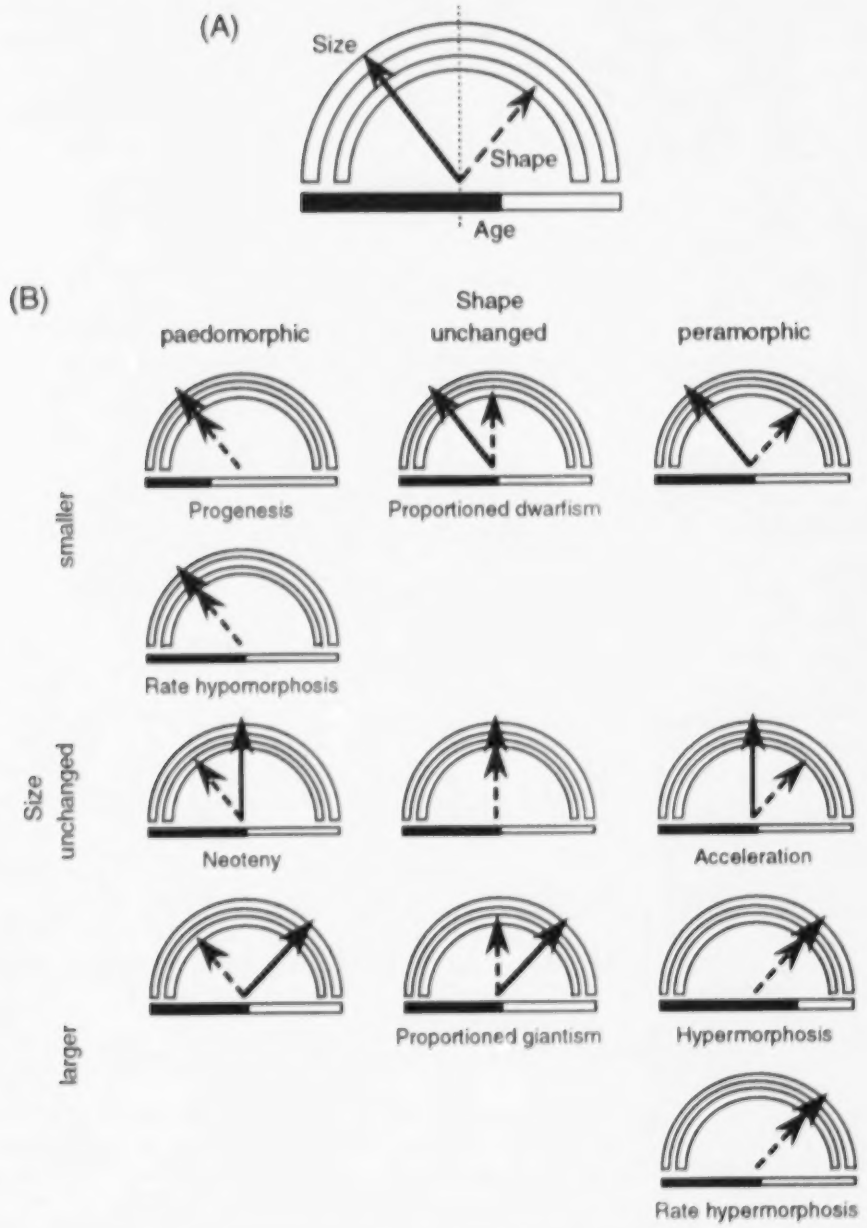


Fig. 3-2. The clock model of Gould (1977). (A) Explanation of the clock. The solid arrow and the outer scale indicate size, the dashed arrow and the inner scale pertain to shape. The shaded area in the horizontal bar is the marker of age. The vertical dotted line shows the calibration for ancestral size, shape, and age at the standard stage. (B) The classification of heterochronic changes according to Gould (1977), with the additional types (rate hypomorphosis and rate hypermorphosis) proposed by Shea (1983a).

As a starting point for using the clock model, the investigator must define measures of size and shape. Size measures can be single measurements, like body length or weight, or composite measures such as first principal component scores. Angles or ratios of lengths can serve to quantify shape, and multivariate techniques offer a variety of composite shape measures (the section on allometry, below, contains more details on the analysis of size and shape). A particular developmental stage is chosen as a standard for comparison; traditionally, sexual maturity has been taken as a standard, but any other stage can be used, provided it is defined by a criterion other than the size or shape variables used in the analysis. The scales are set so that the size, shape, and age of the ancestor at the standard stage are in the midline of the clock (dotted line in Fig. 3-2A). Then, the scales of age and size at earlier stages are calibrated by interpolation between the initial age and size and the standard stage, and by extrapolating beyond that stage. The scale of shape, however, is calibrated so that the shape values are at the same place on the scale as the size values at the corresponding age. As a result, the hands of size and shape will move together when the clock is run for the ancestor (note that this way to calibrate the shape scale implies that growth must not be isometric in the ancestor).

The descendant ontogeny can now be displayed on the clock. Because the age scale is calibrated with a measure of physical time, the descendant's marker for age moves together with that of the ancestor. In contrast, the positions of the hands for size and shape in the descendant may differ from those in the ancestor at any stage, reflecting evolutionary changes of ontogeny. Moreover, the descendant's hands of size and shape may not move together, indicating dissociation and consequently a difference between ancestral and descendant allometries.

At any particular stage of ontogeny, the descendant's size may be larger or smaller than that of the ancestor at the corresponding stage. The position of the descendant's hand of shape indicates if it is paedomorphic or peramorphic—this is a graphical presentation of a shape comparison at the standard stage and with regard to the particular shape measure chosen. Furthermore, the descendant may reach the standard stage at a younger or older age than the ancestor; although the age scale is calibrated by physical time, the intrinsic time scale may change from ancestor to descendant, and the descendant's age marker may therefore be to the left or right of the clock's midline at the standard stage.

The classification of heterochrony in the framework of the clock model is based on the positions of the clock's hands and the age indicator when the descendant reaches the standard stage (Fig. 3-2B). Gould (1977) named six basic types of changes, partly adapting them from the scheme proposed by de Beer (1958), and Shea (1983a) added two further types. *Neoteny* and *acceleration* are changes in shape only, which do not affect age or size at the standard stage. They result in an altered allometry. Early or delayed termination of the descendant ontogenies, while both size and shape retain the ancestral growth rates, result in *progenesis* and *hypermorphosis*, respectively ("time hypomorphosis" and "time hypermorphosis" in the terminology of Shea 1983a). Shea (1983a) proposed the terms *rate hypomorphosis* and *rate hypermorphosis* for changes in the growth rates of size and shape, rather than in the age at the standard stage. Shea coined these terms for cases of ontogenetic scaling, where the descendant retains the ancestral relationship between size and shape; the only difference in allometric plots between ancestors and descendants is therefore that

growth trajectories are either truncated or extended. Finally, *proportioned dwarfism* and *proportioned giantism* result from changes in size, but affect neither shape nor age at the standard stage; yet they produce changes in allometries.

As Figure 3-2B shows, the classification does not include all the possible outcomes—not all of them are named (although one of those shown, in the center of figure, represents the case of no evolutionary change). Moreover, this figure does not even contain all the combinations of age with size and shape changes (there are 26 possible combinations involving some change, 18 of which involve shape change). It is therefore clear that the classification cannot appropriately describe all possible outcomes of the evolution of ontogenies. Additional patterns have to be expected. For instance, Gould (1977: Fig. 40) presented a separate clock model for human heterochrony, which does not correspond to any of the "pure" types (nevertheless, he called it "human neoteny" in the figure caption). Furthermore, Shea (1983b) pointed out that there may not be a "global" heterochronic change affecting organisms in a uniform way. He argued that the results of a comparison strongly depend on the structures being investigated. In his example, comparisons between common and pigmy chimpanzees, application of the clock model yields different results in separate analyses for the skull, trunk and fore limbs, and hind limbs.

Gould's clock model serves as a tool to compare ontogenies and it is the basis for a classification of evolutionary changes in ontogenies; it is used in two rather different ways for these purposes. As a device to display and compare the ontogenies, the model emphasizes the *processes* that produce evolutionary change, that is, evolutionary changes in the dynamics of growth as visualized by the moving hands of the clock. The clock model has rarely been used in this context—computer animation may be more suitable for this purpose than the printed page (a series of clocks at successive stages might serve the same purpose). It is important to note, however, that the classification built on the clock model is based on the results, or *pattern* generated by those processes, because the size, shape and age of ancestor and descendant are compared exclusively at a single standard stage. In this context, only the static configuration of the clock's hands is considered; the same configuration might result despite differences in the developmental dynamics during the ontogenetic stages preceding the one at which the comparison is made.

The Formalism Based on Growth Functions

Because heterochrony deals with changes in rates and timing of growth processes, the most straightforward way to study it is to compare the actual curves depicting measures of size or shape as a function of developmental time. This approach was chosen by Alberch et al. (1979), who based their formalism for the analysis of heterochrony on a simple descriptive model of a growth process. In their model, the growth curve is determined by three parameters: the *time of onset* (α), the *growth rate* (k), and the *time of termination* of growth (β, offset time; although they defined this last parameter as "either a specific age, or a limiting size or shape" [p. 301], it is preferable to use only the time of termination, because the model does not produce the predicted changes if a fixed limit for the growth variable itself is used). An evolutionary alteration in any of these parameters constitutes a heterochronic change. The fourth parame-

ter in the model, initial trait value (y_0), is the result of development before observations are made; changes in this parameter are not directly relevant to the analysis of heterochrony if the study is limited to a well-defined ontogenetic period. In a more mechanistic developmental context, however, evolutionary changes in initial size and shape may be important, as they may influence subsequent growth processes (e.g., Oster et al. 1988; Atchley and Hall 1991). In the phenomenological model discussed here, such changes would be interpreted as heterochronic changes affecting early ontogeny, without explicit reference to their mechanistic cause.

Alberch et al. (1979) adapted the classification scheme of Gould (1977) to the new framework, including the separation of size and shape. Unlike the classification based on the clock model, however, the categories of heterochronic changes are defined by the way they affect the dynamics of growth. The question if a descendant is paedomorphic or peramorphic relative to a given ancestor is therefore separate from the question of which differences in their ontogenies caused this outcome. Rather than considering the morphological results of growth in ancestors and descendants compared at a standard ontogenetic stage, the formalism of Alberch et al. deals with evolutionary modifications of growth itself. The three parameters of the model used to describe growth curves provide the basis to compare the growth dynamics of ancestors and descendants. Alberch et al. (1979) recognized that changes in single parameters can produce the heterochronic changes named by Gould (1977) for the clock model, and therefore they applied the same terms (pp 304–306). Yet, because the clock model compares the results of ontogenetic change, this correspondence is not perfect, especially with respect to changes in the onset parameter, which is not included in the clock model.

An increase in the rate of development for shape corresponds to acceleration, a decrease is neoteny. Proportional dwarfism and giantism are characterized by a lower or higher growth rate for size, respectively. The original version of the formalism assumes the times of onset and termination of development to be the same for size and shape. Earlier or delayed termination in the descendant correspond to progenesis and hypermorphosis, respectively. It is important to note that the definitions provided by Alberch et al. did not relate cessation of somatic growth and sexual maturation in any specific way (although the examples in their paper do); termination of growth can be at any stage and can be independent of sexual maturity (Alberch et al. 1979:302). This is a marked difference to the concept of de Beer (1958). To accommodate changes in the third parameter, onset time, Alberch et al. (1979) coined the terms *predisplacement* and *postdisplacement* for early and delayed onset of development.

As in the clock model, Alberch et al. (1979) used special terms for dealing with rate changes for size (proportioned dwarfism and giantism) and shape (neoteny and acceleration). This separation of size and shape has been abandoned by many recent authors, who have applied the terms originally devised for shape to size data as well (e.g., Creighton and Strauss 1986; McKinney 1986, 1988; McKinney and McNamara 1991; Klingenberg and Spence 1993; Ravosa et al. 1993; Vrba 1994; McKinney and Gittleman 1995). Both size and shape can be used as measures for the "degree of development" of ancestors and descendants, because both are intimately linked by ontogenetic allometry in most organisms (other authors disagree with this reasoning, insisting that

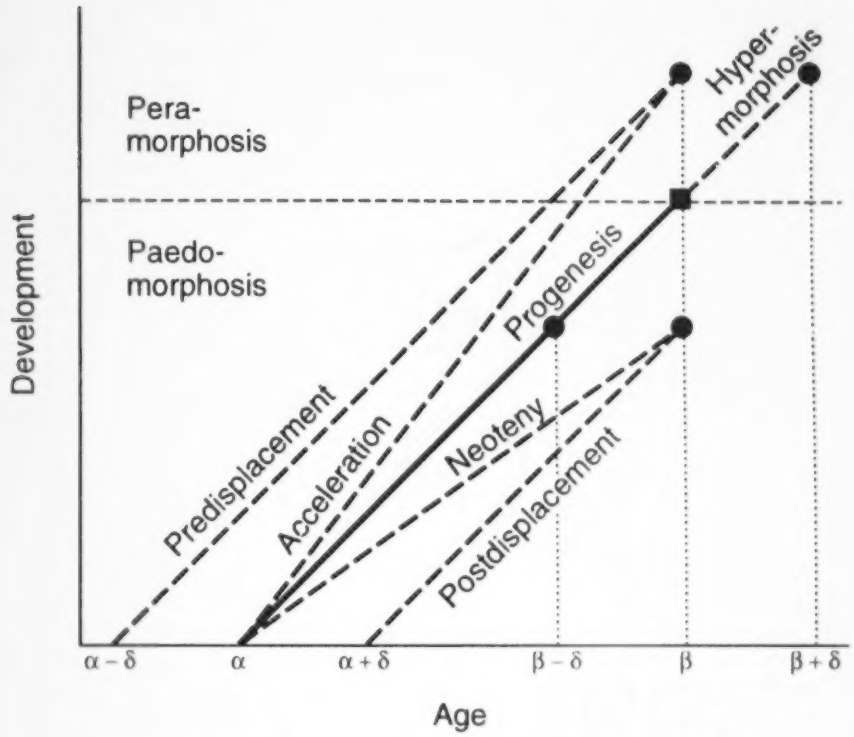


Fig. 3-3. The formalism of Alberch et al. (1979). A measure of shape or of size (only in the modified version of the formalism; see text) is graphed on the vertical axis to indicate the degree of development. The solid line represents the growth trajectory of the ancestor, and the square its morphology at the termination of growth; corresponding symbols for descendants are dashed lines and dots. (From Klingenberg and Spence 1993, with permission of the Society for the Study of Evolution.)

development refers only to shape defined as ratios of measurements; see Godfrey and Sutherland 1995a, 1995b, 1996). Shape results from the relative sizes of an organism's parts; the changes in developmental processes that determine the sizes of organs and of the whole organism therefore are also the changes that affect shape. Therefore, it is logical to apply the same formalism for heterochrony, with separate initial values, rates, as well as onset and termination times for each trait (Fig. 3-3). This is consistent with the three-parameter model for growth curves, which Alberch et al. (1979:301) explicitly proposed for either size or shape. As a consequence of this shift in definition, the terms that originally were used for heterochrony of shape now are applied to both size and shape measures, and the terms "proportional dwarfism" and "proportional giantism" are superfluous in the modified formalism. On the other hand, it is even more important that investigators specify clearly which traits they consider in this expanded framework for analyzing heterochrony.

Because it is based entirely on the simplified model of a developmental process, the formalism of Alberch et al. (1979) does not make any reference to sexual maturation, which was the frame of reference in de Beer's (1958) discussion of heterochrony; ancestors and descendants can be compared at any corresponding stage in their ontogeny (choice of this stage determines the parameter β). Despite the fact that there is no necessary connection, studies that have applied the heterochronic concepts have focused almost exclusively on late ontogeny, where the cessation of development coincides with reproductive maturity, or may even be causally related to it. Several recent critiques of the classical frameworks have targeted this connection to reproductive maturation (e.g., Raff and Wray 1989; Reilly et al. 1996). It is therefore imperative that the frame of reference be made clear in each study; the use of a purely descriptive terminology avoiding the traditional terms is a possible alternative (see Raff and Wray 1989).

Heterochronic processes can be combined, as more than one of the parameters of the growth function can change simultaneously (see also Fig. 2 of Reilly et al. 1996). Only the combinations of heterochronic processes that affect the same parameter in opposite directions are impossible. In pairwise combinations, heterochronic processes either tend to reinforce or compensate the morphological effects of one another (Fig. 3-4). Dommergues et al. (1986) considered pairwise combinations of heterochronic processes, and also presented an elaborate terminology for the resulting heterochronies. In contrast to the results in Fig. 3-4, Dommergues et al. (1986) considered neoteny to be incompatible with predisplacement, as well as acceleration with postdisplacement. Presumably, they regarded these combinations as incompatible because perfect compensation of morphological effects may occur, in which case *no* evolutionary modification will result in either age or morphology at the termination of growth. As Fig. 3-4 shows, however, both these combinations have an effect on the growth trajectories, although the resulting adult morphology may be the same ("isomorphosis" of Reilly et al. 1996). This again illustrates the importance of considering the dynamics of growth processes in detail. Combinations of heterochronic changes, rather than "pure" processes, are to be expected in nature (e.g., Alberch et al. 1979:307); such combinations have been found in several comparative studies of growth dynamics (e.g., Creighton and Strauss 1986; Wayne 1986b; Ishikawa and Namikawa 1987; Klingenberg and Spence 1993; Leigh and Shea 1996).

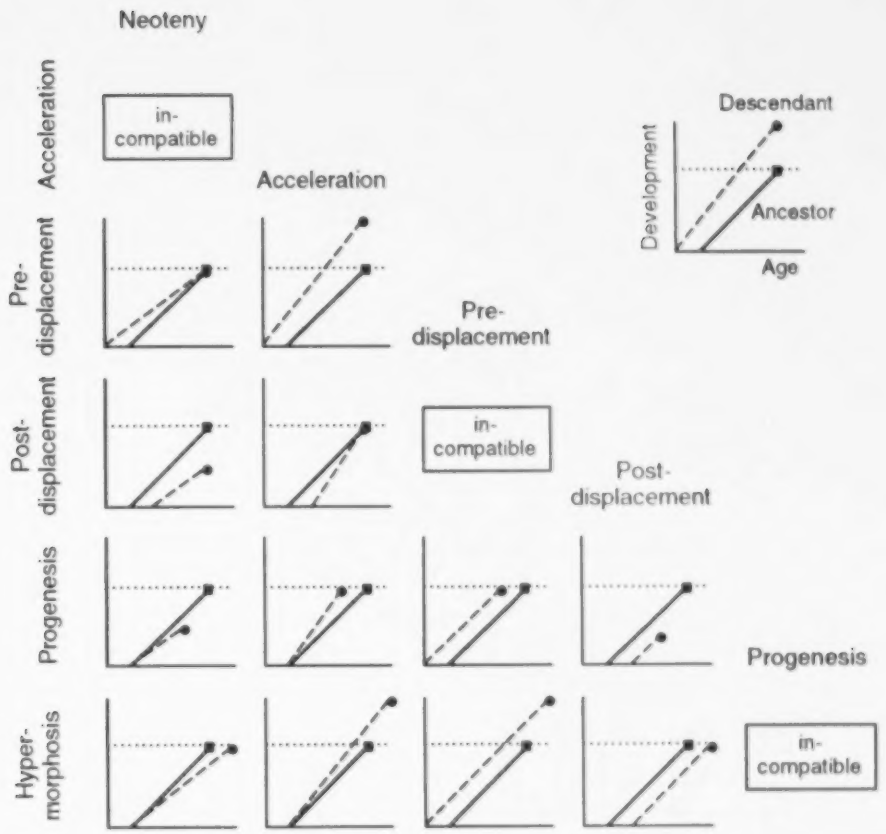


Fig. 3-4. Effects of pairwise combinations of heterochronic processes. For simplicity, the morphological effects of each heterochronic process have been set to a constant value, as in Fig. 3-3. Therefore, the effects either double or cancel out completely, depending on the combination.

The growth model underlying the formalism for analyzing heterochronic changes is a drastic simplification of growth dynamics. Therefore, a crucial step for the application of this framework is the translation from the complex, nonlinear growth functions of real organisms to the simple changes considered in the model (see also the discussion of this "parameterization problem" by Atchley 1987). The rate parameter is often derived from the average rate of growth: the total growth increment divided by the time between the onset and termination of growth. As an alternative, however, the growth rate at a particular stage or the maximal growth rate can be used. The age of onset and cessation of growth are often difficult to determine, because size or shape measures gradually reach an asymptote. As a proxy, investigators may choose the age at which the variable reaches a certain percentage of the asymptotic value or when the growth rate exceeds a particular threshold value or a given fraction of the maximal rate (Fig. 3-5). These choices can affect the results of the analysis, and proper care is necessary for interpreting the results.

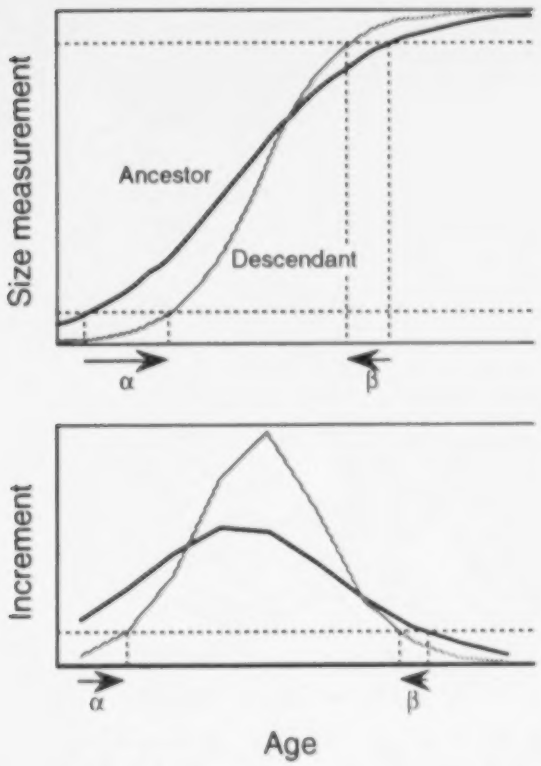


Fig. 3-5. Comparison of growth dynamics in two species. The top panel shows the growth curves, and the bottom panel the growth increments at regular intervals (e.g., annual growth). Each panel illustrates a different way to determine the times of onset (α) and cessation (β) of growth: in the top panel, α and β are defined as the times at which size reaches a given percentage of the final value, and in the bottom panel as the times at which the growth proceeds at a certain threshold rate. The results may differ depending on the method and threshold values chosen, but in this case, they are consistent: the descendant starts growing later, grows at a higher rate (both average and maximum), and ceases to grow earlier than the ancestor. Therefore, the heterochronic change is a combination of postdisplacement, acceleration, and progenesis.

An example of how this formalism can be applied to real organisms is the study by Creighton and Strauss (1986), comparing growth among several species of rodents. These authors used von Bertalanffy curves (a function with exponential decay of growth rate) to quantify the growth of each species. They defined β, the offset time parameter, as the time at which the growth curve attained 90% of the asymptotic size. The average growth rate between birth and the offset time β, could be used to identify neoteny and acceleration. Creighton and Strauss took birth weight as an indicator for prenatal development, although this criterion cannot directly recognize pre- or postdisplacement. Comparing parameters of growth functions fitted to the data requires that the shape of the curves is constant across the study group, that is, that the underlying

model (e.g. von Bertalanffy, Gompertz) accurately represents the growth curves of all species. Other studies focused on direct comparisons of growth curves, without applying a particular growth model (e.g., Ishikawa and Namikawa 1987; Strathmann et al. 1992; Klingenberg and Spence 1993). In addition, nonparametric regression techniques can be very useful for computing both cumulative growth and velocity curves (e.g., Guihard-Costa 1991; Leigh and Shea 1996). The chief advantage of nonparametric techniques is their flexibility, because they do not assume a particular shape of the growth curves (for a comparison of parametric and nonparametric methods, see Leigh and Shea 1996).

Different structures have their separate ontogenies, which can be described by their own growth trajectories. Despite the overall integration of ontogenies, growth in various organs can be controlled intrinsically (e.g. Bryant and Simpson 1984); this provides different structures with a degree of independent variation, and therefore they may also evolve in different ways. In the analysis each organ has its own set of parameters for onset, rate, and termination of growth (Atchley 1987; Atchley and Hall 1991). Whereas the original formalism of Alberch et al. (1979) assumed that the times of onset and termination of development were the same for size and shape ("global heterochrony", McKinney and McNamara 1991), the modified formalism allows separate heterochronies for each character ("mosaic heterochrony," David 1989, 1990; "dissociated heterochrony," McKinney and McNamara 1991).

A clear example of separate growth trajectories is growth in birds: whereas leg measurements reach an asymptotic value relatively early, the wings continue to grow longer (Boag 1984; Carrier and Leon 1990), and the bill may only reach its final size several weeks after fledging (Boag 1984). As a result, allometries may change from one growth phase to another (e.g., Cane 1993), and different organ systems may evolve opposite heterochronies, such as in the Galápagos cormorant and other flightless birds, which have a paedomorphic pectoral girdle and a peramorphic pelvic girdle relative to their hypothetical ancestors (Livezey 1992, 1995).

Combinations of heterochronic processes may be of crucial importance for ecological interpretations of heterochrony. Rate and timing of development are important determinants of life history, which is the interface between an organism's ontogeny and its environment. Gould (1977) argued that heterochrony, through its connection to life history parameters, may be correlated with the dynamics of populations and their environments. Specifically, he hypothesized (pp 290–294) that neoteny should be associated with *K* selection and progenesis with *r* selection, respectively (see also McKinney and Gittleman 1995). In partial support of the hypothesis, McKinney (1984, 1986) found an association between the size of fossil echinoids and the stability of their habitats. Other studies also reported associations between environmental conditions and evolution by heterochrony (Allmon 1994; Wei 1994). Due to the absence of age data, however, it is not possible to interpret these results in terms of heterochrony with any degree of confidence; size may be a misleading proxy for age (see below; Klingenberg and Spence 1993; Godfrey and Sutherland 1995a).

Conflicts Between Different Frameworks

Although the classifications of heterochronic processes based on both Gould's (1977) clock model and the formalism of Alberch et al. (1979) mostly use the same terms, there is no strict one-to-one correspondence between them, and their application may sometimes lead to contradictory conclusions. Both frameworks adapted some of de Beer's (1958) categories, and therefore were designed to capture the essence of his notions, as well as to be faithful to the original examples for which the terms were coined (for discussion, see Gould 1977). There is a substantial difference, however, in the way the two formalisms identify categories of heterochronic changes: the clock model makes a static comparison of size, shape, and age at a particular stage chosen as a standard, whereas the formalism of Alberch et al. (1979) compares specific parameters of growth functions. The differences have become even more important since most authors have abandoned the distinction between size and shape as the measure for the degree of development (Fig. 3-2; McKinney 1988; McKinney and McNamara 1991; Klingenberg and Spence 1993; Vrba 1994).

To illustrate these differences, I compare the two terminologies by simply describing the categories of the clock model (Fig. 3-2) with the terms used in the modified version of the formalism by Alberch et al. (Fig. 3-3). The clock model's progenesis is also termed progenesis in the latter formalism, but this must be specified for both size and shape separately. Rate hypomorphosis, however, is described as neoteny of both size and shape (or as neoteny for shape, combined with dwarfism). For neoteny and acceleration of the clock model, either the corresponding processes or, alternatively, postdisplacement and predisplacement act on shape; there is no change in size in either case. Conversely, proportioned dwarfism and giantism are described as neoteny and acceleration in size, with no change in shape. The clock model's hypermorphosis is characterized as hypermorphosis in both size and shape, whereas rate hypermorphosis is acceleration of both size and shape.

Attempts to infer the types of heterochrony from plots of growth functions illustrate the ambiguities that can be produced by switching between the two systems without proper caution. For instance, Richtsmeier and Lele (1993: Fig 13a) plotted a growth variable $G(t)$, for example a linear distance measurement, as a function of time; they also plotted the specific growth rate, the derivative of the log-transformed measurement with respect to time. The resulting plots, although intended to depict rate hypermorphosis, are indistinguishable from plots for acceleration in the framework of growth functions introduced by Alberch et al. (1979). As all the heterochronic types in the clock model, rate hypermorphosis can only be identified by considering size, shape, and time simultaneously (Fig. 3-2). In order to generate rate hypermorphosis of the clock model, the relative increase of specific growth rates in all traits must be the same, so that the descendant ontogeny follows the allometric trajectory of the ancestor. This example demonstrates that concepts and terminology of the two frameworks are not directly compatible: because the clock model simultaneously refers to both size and shape, the heterochrony types of the clock model cannot be read from a single growth function.

An Example: Human Heterochrony

The debate on heterochrony in human evolution is a particularly clear example of how conflicting interpretations may arise because of the differences in the concepts on which analyses are based, even if all these analyses are carried out correctly. In the last few decades, the hypothesis that humans are neotenous relative to their ancestors has dominated this debate (de Beer 1958; Gould 1977; Montagu 1989). Recently, however, this view has been challenged, most notably by Shea (1989), who presented a detailed critique of the neoteny hypothesis and concluded that no single heterochronic process dominated human evolution (see also Dean and Wood 1984; Wood 1996). McKinney and McNamara (1991) also criticized the neoteny hypothesis, but they argued that hypermorphosis was the dominant process instead. Whereas Shea (1989) agreed with earlier authors that some morphological features of humans are paedomorphic, although not by neoteny, McKinney and McNamara (1991) interpreted most of these features as peramorphic. In contrast, Vrba (1994) explained the evolution of the increased brain size and numerous paedomorphic features through prolongation of biphasic growth. Godfrey and Sutherland (1995a, 1995b, 1996) criticized several of the studies arguing against the neoteny hypothesis on methodological grounds. Because these studies used the entire array of conceptual frameworks to analyze the same problem, a closer examination of the arguments made by different authors can serve as a case study of the application of heterochronic formalisms and of their modifications.

Like earlier authors, de Beer (1958:68–76) based his argument in favor of neoteny as the dominant process in human evolution on an enumeration of traits consistent with this pattern. First, he reviewed evidence supporting paedomorphosis of various human features (pp 68–73) and then listed delays of multiple developmental events in humans relative to the great apes and other primates (pp 73–76). This reflects de Beer's notion of neoteny as paedomorphosis through retardation of development (1958:36, 63 ff.). Montagu (1989) took the enumerative approach farther; most of this book is devoted to listing purportedly paedomorphic characteristics of humans (Montagu treated neoteny and paedomorphosis as synonyms). Montagu applied the concept of paedomorphosis not only to physical traits, but among others also included love, friendship, sensitivity, work, optimism, honesty, song, and dance, all of which he considered neotenous "[b]ecause they appear so early in life of the child, and many of them are already present during fetal life" (p 107). It is unclear, at the very least, what criteria Montagu applied to designate these properties as paedomorphic (see also Shea 1989:94; McKinney and McNamara 1991:309 f.).

Gould's (1977) argument for human neoteny is of a different form, as he emphasized the weakness of the enumerative approach (pp 363 ff.; but see also Shea 1989:88 ff.). He first noted that general temporal retardation of development characterized human evolution, an observation uncontested even by critics of the neoteny hypothesis, and then asserted that "retardation established a matrix within which all trends in the evolution of human morphology must be assessed" (p 365). "Retardation" is at the heart of the discussion about the evolution of human ontogeny, but this term is responsible for much of the confusion. The word "retardation" can denote a slowing of a continuous process or the delay of discrete events; it can thus refer to either rates or timing. Neither

de Beer (1958) nor Gould (1977) specified in which sense they used the word; in this review, I strictly use it to mean a delay in the timing of an event. Many of the events affected by retardation relate to maturation of the reproductive system (e.g., puberty and the adolescent growth spurt). Retardation, in this sense, does not automatically imply paedomorphosis (e.g., Gould 1977: footnote on p. 376), which is essential for a heterochronic change to be recognized as neoteny under the clock model; instead, paedomorphosis must be established as a separate fact. Stated in the terms of the framework of growth functions, retardation only produces paedomorphosis if the slowing of development is strong enough to outweigh the effects of prolonged development time in the combination of neoteny with hypermorphosis (see Fig. 3-4).

A "matrix of retardation" was the principal factor in Gould's (1977) account of paedomorphosis pervading human evolution, because "[g]eneral retardation of this sort entails extensive paedomorphosis as an almost ineluctable consequence" (footnote on p. 376). This "matrix" implies general paedomorphosis that has a common developmental basis, while individual traits may deviate for specific reasons. For instance, Gould used the "matrix of retardation" in his arguments to counter claims by critics of the neoteny hypothesis that some human features were peramorphic, such as the chin and prominence of the nose. Most of these arguments focused on the definition of the "shape" of these features relative to neighboring structures: Gould suggested that these traits only appear to differ from other parts of the skull because of the more pronounced paedomorphosis of surrounding tissues by extreme retardation (pp 380–382). He acknowledged, however, that there are exceptions to human neoteny, for example the legs that evolved by hypermorphosis—as another facet of general retardation.

Shea (1989) evaluated the hypothesis of human neoteny by comparing the predictions of the clock model to the data that had previously been used to support the hypothesis. According to the clock model (Fig. 3-2), the neoteny hypothesis predicts that shape in human adults, the standard stage chosen, should correspond to the shape of ancestral juvenile stages; changes in age or size are in addition to those predicted by "pure" neoteny (compare Gould's [1977] separate clocks for neoteny [his Fig. 39B] and "human neoteny" [his Fig. 40]). Shea (1989) pointed out that many of the paedomorphic features previously cited in support of the neoteny hypothesis result from rate hypomorphosis rather than neoteny of the clock model, and are associated with allometric scaling. Note, however, that in the modified formalism of growth functions, rate hypomorphosis would be described as neoteny of both size and shape, thus altering the implications of the neoteny hypothesis.

Because the neoteny hypothesis under the clock model predicts change in shape without a corresponding change in size, Shea (1989) pointed out that allometric relationships should differ between ancestors and modern humans under this hypothesis. He discussed intraspecific variation in humans cited previously as examples for neoteny, namely sexual dimorphism and the growth of pygmies, and pointed out that this variation mainly concerned the extent of growth along a shared allometric trajectory (pp 76–80). In addition, ontogenetic scaling is generally widespread in primate evolution (see, e.g., Shea 1983a, 1992a, 1992b; Ravosa et al. 1993). Godfrey and Sutherland (1995b:422) criticized Shea's analyses because a conserved allometric trajectory for one shape variable does not rule out dissociation for other shape vari-

ables. Dissociation in any single shape variable automatically implies dissociation for "shape" in the multidimensional space of all shape variables, but other shape variables may retain their ancestral association with size. The fact that some shape variables therefore may "miss" dissociation is a shortcoming of all heterochronic analyses, because both the clock model and the formalism of growth functions must treat the "shape" or "development" variable and its rate of change as scalar values. The choice of this variable is a crucial step in the analysis of heterochrony. Separate analyses of various shape variables, each emphasizing different aspects of shape, may well lead do opposite results; these are only apparently contradictory, but point out the complexity of the observed changes. Multivariate analyses are a possible alternative, but in turn, their findings cannot always be interpreted correctly in the simple one-dimensional perspective of paedomorphosis versus peramorphosis.

Shea (1989:80–85) also challenged Gould's (1977) explanation of human paedomorphosis by the "matrix of retardation" because of the lack of empirical evidence supporting an association of extended development with paedomorphosis, either among human populations or among species of primates or other animals. There clearly must be an association between rate of development and the time at which it reaches a given threshold value (Godfrey and Sutherland 1996:36 f.), but beyond this, the implications of developmental delay are open. In a more general context, it is important to recognize that retardation has two opposite facets, depending on whether the emphasis is on the stage before or after a particular developmental event. On the one hand, organisms with retarded ontogenies may have a more paedomorphic appearance because the effects of delayed developmental events have not yet appeared at the ages when comparisons are made. On the other hand, however, developmental processes that take place before this event have more time to accumulate a stronger effect by hypermorphosis, which is peramorphic with respect to the polarity for this ontogenetic stage.

Although seemingly self-evident, this perspective may be helpful to reconcile the contrasting positions on the development and evolution of the human brain size. Ontogenetic allometries of brain size versus body size change drastically during development, and as a consequence, the polarity of paedomorphosis versus peramorphosis reverses itself. During fetal development of humans and other mammals, brain size increases fast and with positive allometry (although slight) relative to total body size, but there is a switch to a slower rate and negative allometry at birth or some time thereafter; this switch occurs especially late in humans (e.g., Gould 1977:371–373; Shea 1989:82; McKinney and McNamara 1991:301–303). This prolongation of brain growth at the high fetal rate is responsible for a marked increase in relative brain size, and it makes modern humans clearly peramorphic by hypermorphosis relative to their ancestors, if the ontogenetic polarity of fetal development is used as the base of the comparison (e.g., McKinney and McNamara 1991). In contrast, the same change renders humans paedomorphic if the ontogenetic polarity of the later postnatal period is applied (Gould 1977; Shea 1989). Awareness of the reversal of ontogenetic polarity can help to resolve an apparent paradox, as illustrated by the following quotation: "Our relatively large brains therefore result from *time hypermorphosis* [i.e., they are peramorphic in the fetal polarity], and they yield a high brain/body ratio that is paedomorphic, given the general postnatal negative allometry of brain/body growth" (Shea 1988:252 f.).

A similar logic is behind the terms "hyper-paedomorphosis" and "hypo-peramorphosis" coined by Vrba (1994), which also pertain to biphasic (or multiphasic) growth. In the model of "hyper-paedomorphosis" a period of fast development is followed by a period with a slower rate. An extension of both periods by an equal proportion leads to an increase in the developmental change achieved in both periods, and is thus peramorphosis by hypermorphosis. Because the first period has a higher rate, however, the increase of developmental change during this period is larger than in the second period, and a higher proportion of the total ontogenetic change stems from the first period in the descendant than in the ancestor. In an apparent redefinition of the term paedomorphosis, Vrba (1994) concludes that "the proportion of shape units derived from the earlier juvenile phase (or paedomorphic shape) increases in the descendant" (p 359). This concept of paedomorphosis and peramorphosis, according to the time when shape change occurs, is fundamentally different from the concept used by other authors (although remarks of Gould [1975:286; 1977: footnote on p 365] and Shea [1989:82] may be interpreted to foreshadow Vrba's redefinition) and it is even inconsistent with the definition in the glossary of Vrba's chapter (p 371). In all of Vrba's models, both developmental periods have the same polarity (a monotonic increase of the shape score), and therefore any extension of growth periods leads to peramorphosis, whereas truncation produces paedomorphosis.

McKinney and McNamara (1991) analyzed human heterochrony with the modified version of the formalism by Alberch et al. (1979); much of their argument therefore concerns size, and not exclusively shape (but see the critique by Godfrey and Sutherland 1995b, 1996). They confirmed the pervasive retardation of human ontogeny, but they argued that rates of developmental processes in humans are not lowered, and human evolution is therefore dominated by peramorphosis through the process of hypermorphosis, not paedomorphosis by neoteny. This reasoning required a reversal of the ontogenetic polarity previously hypothesized for many morphological traits. In this context, McKinney and McNamara (1991:292) raised a fundamental criticism against the use of shape measures, echoing the more extensive analysis of Shea (1989:85–88). For instance, they argued that the similarity between the rounded skull of adult humans and the fetal or juvenile skull of other hominoid primates is a consequence of the prolonged retention of high fetal growth rates of the brain in humans, and therefore a hypermorphic character. Tension and pressure by the expanding brain are important factors in the growth of the braincase (Herring 1993:176 f.), and retention of a rounded skull may have been selected for as an efficient way for accommodating an enlarged brain while maintaining other functions of the head (Ross and Henneberg 1995). Therefore, McKinney and McNamara (1991) considered the slowing-down of the development of overall shape to be only an apparent by-product of this process, which does not reflect a slowing of underlying growth processes. In sum, Shea (1989) as well as McKinney and McNamara (1991) pointed out that biological processes should be the principal considerations in analyses of heterochrony, and that similarity exclusively based on a geometric definition of shape may be superficial.

Sexual dimorphism traditionally has been used as an argument for human neoteny. For example, Montagu (1989:27) stated that "[t]he female skull ... is more pedomorphic in all human populations than the male skull; this holds true for many other somatic traits and, I have not the least doubt, for functional and

behavioral traits as well." In contrast, Shea (1989:84) and McKinney and McNamara (1991) noted that sexual dimorphism does not conform to the predictions for neoteny. McKinney and McNamara argued that females are progenetic relative to males because growth rates are about the same in both sexes, but the adolescent growth spurt and termination of growth occur about two years earlier in girls than in boys (Marshall and Tanner 1986; Tanner 1989). In the African apes, however, sex differences in growth curves lead to size dimorphism in a variety of ways and may be subject to natural selection themselves (Leigh and Shea 1996), suggesting that there is considerable evolutionary flexibility in the way sexual dimorphism develops. Such flexibility raises the question of just how informative sexual dimorphism is for understanding how humans evolved as a species.

The discussion about human heterochrony has been dominated by disagreement about interpretations of the same facts. Retardation, in the sense of a prolongation of the entire ontogeny, has been accepted by most workers in the field (e.g., Gould 1977; Montagu 1989; Shea 1989; McKinney and McNamara 1991), and has been confirmed recently with new methods applicable to fossil hominids (see Smith and Tompkins 1995). The disagreements about human heterochrony largely stem from differences in the definition of traits to be analyzed, not from factual differences or errors in the analyses. Disparate concepts of "size" and "shape" (see the section on allometry, below) and differing formalisms for analyzing heterochrony have led to conflicting interpretations. Generalizations about evolutionary process on the basis of published results are currently impossible because of these discrepancies—one author's "neoteny" is another author's "hypermorphosis." These disagreements about underlying frameworks are deeply entrenched, and there is little or no prospect of a uniform terminology anytime in the near future. The only remedy for this situation is both for authors to be explicit about the criteria on which diagnoses of heterochronic processes are based, and for readers to pay close attention to these issues when comparing the results of different studies.

Formalisms Based on Developmental Processes

The methods for analyzing heterochrony presented in the preceding sections are entirely phenomenological. They examine the evolution of growth at the scale of the whole organism, characterizing them by growth functions derived from external measurements. Such growth curves reflect the aggregate dynamics of a multitude of unknown developmental mechanisms that work at the organismal, tissue, cellular, and molecular scales, and their interactions. Moreover, most studies of heterochrony deal with late stages of ontogeny.

Modern developmental biology, conversely, is concerned mostly with early stages and smaller, more localized scales. Pattern formation and cell differentiation are the main focus, rather than growth. Developmental biologists mainly focus on individual processes, which are often known in great detail, but usually for only a few experimental organisms. In recent years, interest in evolutionary problems has increased among developmental biologists, and a growing number of studies have discussed heterochrony in relation to its mechanistic basis (e.g., Hall 1984, 1992; Ambros and Horvitz 1984; Alberch 1985; Lord and Hill 1987; Ambros 1988; Parks et al. 1988; Regier and Vlahos 1988; Raff and Wray 1989; Wray and Raff 1990; Raff 1992; Duboule 1994; Richardson 1995; Gilbert et al. 1996).

Cell proliferation and morphogenetic movements can be described by rates and timing, and to some extent even processes such as gene regulation and other molecular interactions. Therefore, evolutionary changes in these parameters can be interpreted as heterochrony in a manner analogous to the frameworks discussed above. Raff and Wray (1989) proposed an alternative classification of heterochrony, which is similar to the formalism of Alberch et al. (1979), but Raff and Wray specifically considered the regulatory mechanisms that underlie evolutionary changes in rate, onset or termination of a developmental process. Whereas the basic formalism (see Fig. 3-3) describes developmental phenomena exclusively as a function of time, Raff and Wray (1989) also include processes whose product determines the timing of termination via a feedback mechanism. For instance, in a such a product-regulated process, an evolutionary increase in the rate will cause the process to terminate early in a descendant; in the phenomenological framework, this would be described as a combination of acceleration and progenesis (Fig. 3-4). Because the regulatory mechanisms that lead to such combined heterochronies differ from those based on timing only, Raff and Wray (1989) considered the combinations of heterochronic processes that lead to compensation as separate processes. The descriptive terminology for heterochrony proposed by Raff and Wray does not include any of the terms used in the classical frameworks of heterochrony (for a synonymy, see Table 3-1).

This approach can be adapted to particular examples. The "morphogenetic triangle" proposed by Keene (1982, 1991) is a formalism specifically devised for describing tooth growth. Crown height of a tooth is determined by proliferation of the inner enamel epithelium and its subsequent differentiation that initiates mineralization (see also Smith 1995). Proliferation starts earlier and proceeds at a slower rate than differentiation. Growth of the tooth crown is

Table 3-1. Comparison of the terminology for heterochrony proposed by Raff and Wray (1989) with the modified formalism of Alberch et al. (1979; see Figs. 3-3, 3-4).

Raff and Wray (1989: Table 2)	Alberch et al. (1979)
Early initiation (a)	Predisplacement
Early initiation (b)	Predisplacement with progenesis
Late initiation (a)	Postdisplacement
Late initiation (b)	Postdisplacement with hypermorphosis
Early termination	Progenesis
Late termination	Hypermorphosis
Faster rate (a)	Acceleration
Faster rate (b)	Acceleration with progenesis
Slower rate (a)	Neoteny
Slower rate (b)	Neoteny with hypermorphosis

terminated when differentiation catches up with proliferation. Hence, a graph can be set up with two lines representing the progression of proliferation and differentiation as a function of age; these lines and the time axis enclose a triangle whose tip indicates the time of completion and final height of the tooth crown. An increase in height can be achieved by a higher rate or earlier onset of proliferation, or by a lower rate or later onset of differentiation. Although the "morphogenetic triangle" is based on a very simplified model of the mechanisms underlying tooth development, Keene (1991) showed that it can be usefully applied to explore variation in tooth size, shape, and features such as the number of cusps.

Detailed knowledge of developmental mechanisms can lead to a deeper understanding of the potential for evolution by heterochrony. Ontogenies are not merely sequences of events occurring in a fixed temporal order, which can be speeded up or slowed down. To the contrary, there are various kinds of relations among developmental events, and change in any one event may or may not have a cascade of effects on later stages. Two successive events in a developmental sequence may be directly related to one another as cause and effect, they may both be the effects of a third, earlier event, or alternatively, they may not be related, but only occur in sequence by coincidence (Alberch 1985; Raff and Wray 1989). Clearly, the potential for dissociation and heterochrony differs between these alternatives, and therefore developmental causation is of great importance for understanding heterochrony. Nevertheless, there is disagreement on whether heterochronic studies assume that ontogenetic stages are causal sequences of developmental events. In fact, this disagreement is based on fundamentally different views of heterochrony. Alberch (1985) insisted that a causal relationship among events is essential for heterochronic analysis. He referred to the frameworks of Gould (1977) and Alberch et al. (1979), which require size and shape to be monotonic functions of age; these formalisms cannot accommodate changes in the sequence of ontogenetic stages. Therefore, to use the formalism, one must assume the order of stages is invariant; assuming a causal relation of developmental events ensures that the sequence is constant and that stages can be homologized unequivocally. In contrast, Raff and Wray (1989) specifically refer to heterochrony as any change in the relative timing of developmental events, including reversals of the ancestral sequence, and many other developmental biologists follow this definition (e.g., Ambros and Horvitz 1984; Hall 1990, 1992). The absence of causal relations reduces constraints on such changes in the sequence of developmental events, and therefore facilitates evolution by this kind of heterochrony (Raff and Wray 1989).

Developmental biologists tend to define heterochrony as evolutionary shifts in the relative timing between developmental events in an organism, that is, a change in the order of these events (e.g., Raff 1987; Raff and Wray 1989:410, 413; see also Ambros and Horvitz 1984; Hall 1984, 1990, 1992; Wray and McClay 1989; Wray and Raff 1990; Collazo 1994; Duboule 1994; Richardson 1995). This concept of heterochrony is somewhat different from that employed in the frameworks outlined in the previous sections, where heterochrony is defined as any evolutionary change in rates or timing of developmental processes, even if the order of events is unchanged (e.g., McKinney and McNamara 1991). In many studies, this distinction is not stated explicitly, and it appears as a mere shift in emphasis. Still, this shift constitutes a return to the original definition of the term by Haeckel, who defined heterochrony as

changes in the order of appearance of organs (Gould 1977). This definition is more restrictive than the one going back to de Beer, who included the speeding up or slowing down of a conserved ontogenetic sequence (de Beer 1958; Gould 1977, 1992; McKinney and McNamara 1991). This recent trend within modern developmental biology to resurrect Haeckel's original meaning of heterochrony, which Gould (1992:165) thought to be extinct, is an additional twist in the long and changing history of this term (Gould 1977, 1992); it is especially ironic because prominent developmental biologists believe that one of the achievements of Gould's (1977) book was "exorcising the ghost of Haeckel" (Gilbert et al. 1996:363). This reversal in the definition of heterochrony has a parallel in the renewed use of the term heterotopy, coined by Haeckel for evolutionary change in the location from which an organ originates during development (Wray and McClay 1989; Hall 1992:210–212).

Case Studies on the Developmental Basis of Heterochrony

Heterochrony is an evolutionary process that leads to morphological differences between ancestors and descendants by changing developmental pathways (Hall 1992: 199 f.). Understanding the mechanistic basis of these changes and their consequences is a central component of a synthesis between developmental and evolutionary biology (Atkinson 1992; Hall 1992; Raff 1992; Gilbert et al. 1996). A number of model systems have been explored in this context, and developmental biologists increasingly adopt a comparative approach and expand the spectrum of study organisms (but see Hanken 1993). This broader comparative basis and the recent advances in developmental genetics make it possible to pinpoint the actual mechanisms underlying the more traditional models.

A well-studied example of heterochrony in early ontogeny is the evolution of nonfeeding larvae and direct development in sea urchins. Several lineages have evolved development via lecithotrophic larvae with reduced or lost feeding structures as an alternative to their ancestral ontogeny, which includes a planktonic feeding stage, the pluteus larva (e.g. Strathmann 1978, 1985; Raff 1987; Wray and Raff 1991a; Wray 1992). This change in life history is associated with fundamental differences in larval structure and developmental processes, of which many can be interpreted as heterochrony (Wray and Bely 1994; Wray 1995); however, there may not be any corresponding differences in the adult stage. Eggs of species with direct development are much larger and contain more yolk than the eggs of species with feeding larvae (Raff 1987; Wray and Raff 1990, 1991a). Cleavage, blastula formation, and gastrulation differ between the two developmental modes (Raff 1987; Wray and McClay 1989; Wray and Raff 1990, 1991a, 1991b; Wray and Bely 1994). Features characteristic of the pluteus larva in typical sea urchins, such as the larval gut, the calcareous skeleton and arms, are reduced or completely absent in embryos with direct development (Raff 1987). At least some of these changes are related to developmental timing, as larvae of nonfeeding species delay the expression of genes associated with skeletogenesis (Wray and Bely 1994; Wray 1995). The echinus rudiment, however, which consists of structures of the juvenile sea urchin that persist after metamorphosis, develops in a much shorter time than it does in the typical pluteus larva (Wray 1995). Despite these consistent features, there are also substantial differences among the many lineages that acquired direct development independently: developmental pathways dif-

fer between lineages, and even the sequence of the appearance of organs may be reversed (Raff 1987).

Egg size plays a key role in the determination of early ontogeny (Raff 1987; Wray and Raff 1991a; Wray 1995). Sea urchins are a suitable experimental system for manipulating the size of developing embryos because blastomeres separated in the two- or four-cell stage of cleavage develop into complete larvae. Experimental manipulations of egg size can have similar heterochronic effects on larval development as interspecific differences in egg size (Sinervo and McEdward 1988). Increased density of food presented to pluteus larvae leads to reduced growth of larval structures and early metamorphosis, developmental effects which are similar to those of large yolky eggs in sea urchins with direct development (Strathmann et al. 1992). Egg size is not the only factor in the transition to direct development (Wray and Raff 1991a), since there is no pluteus larva after experimental reduction of egg size in direct developers (Okazaki and Dan 1954; Henry and Raff 1990). Nevertheless, an evolutionary increase in egg size could have been a mechanism involved in the origin of nonfeeding larvae, providing a simple developmental basis for multiple heterochronic changes as an evolutionary response to selection on larval life history traits (Sinervo and McEdward 1988; Strathmann et al. 1992). In turn, an increase in egg size can be interpreted as peramorphosis during oogenesis (Wray 1995). Similar associations of changes in egg size, early development, and life history have also been reported in other animal groups (Elinson 1987, 1989; Freeman and Lundelius 1992; Sinervo 1993).

These studies, among others, suggest that relatively simple changes in initial conditions may result in multiple heterochronic changes under a conserved set of developmental rules. This *structuralist* or *internalist* framework for understanding the basis of morphological evolution (Maynard Smith et al. 1985; Wake and Larson 1987; Alberch 1989, 1991) requires knowledge of developmental mechanisms and of the way these mechanisms are integrated.

Shubin and Alberch (1986) and Oster et al. (1988) used the structuralist approach to explain the origin and diversity of vertebrate limb designs (see also Coates 1994; Hinchliffe 1994; Morgan and Tabin 1994; Shubin 1995). The empirical basis of these models comes mostly from interspecific comparisons of development (e.g., Shubin and Alberch 1986; Blanco and Alberch 1992), from the analysis of intrapopulation variation (Shubin et al. 1995), or from the study of teratological "monsters," which provide information about the intrinsic rules of variation precisely because they are maladaptive and often even lethal (Alberch 1989). Intraspecific variants, presumably due to environmental or relatively small genetic differences, can be atavisms or they can correspond to apomorphies of other clades (Shubin et al. 1995), and indicate that the developmental processes channel morphological variation in certain directions, and may thus account for parallel evolution and reversals (Wake 1991).

Insect wings are another well-understood example. In a large comparative study, Nijhout (1990a, 1991) interpreted the diversity of color patterns on the wings of butterflies as different expressions of the same groundplan (see also Nijhout 1994b). Experimental transplantation or cautery of parts of developing wing discs indicated that pattern elements result from the interaction of a inductive focus with the surrounding epithelium (Nijhout 1980, 1985, 1991; French and Brakefield 1992, 1995; Brakefield and French 1995). Moreover, Nijhout (1990a, 1991) developed a reaction-diffusion model, in which an acti-

vator and inhibitor substance interact to establish the pattern in a discrete portion of the developing wing disc. He showed that this model is able to produce nearly the entire diversity of observed wing pattern elements by changing values for thresholds, diffusion constants, and decay rate constants for activator and inhibitor, and the time when the pattern of morphogen concentration is translated into color patterns; most of these changes can be interpreted heterochrony. Subsequently, Carroll et al. (1994) found that the expression of several genes known for their role in pattern formation in *Drosophila* forms spatial patterns in developing butterfly wings that correspond to prospective elements of the color pattern (see also Carroll 1994); this constitutes a possible mechanistic basis for the theoretical model. In a similar way, Garcia-Bellido and de Celis (1992) proposed a model of the development of veins in the wings of *Drosophila*, based on detailed genetic analysis (see also Sturtevant and Bier 1995). The relevance of these studies for evolution is underscored by Powell and DeSalle (1995) who pointed out that several features of wing venation that vary across the family Drosophilidae are morphologically equivalent to mutants known from *Drosophila melanogaster* (phyletic phenocopies).

The preceding examples suggest that at least some evolutionary changes may have a fairly simple developmental basis. Experimental studies showed that factors such as temperature and exposure to certain chemical compounds can produce the phenotypic effects of heterochrony (Yamashita et al. 1991; Blackstone and Buss 1992). Moreover, some studies discovered that mutations at single gene loci may produce heterochronic changes. A clear example are the heterochronic genes in the nematode *Caenorhabditis elegans*, which control the timing of events in postembryonic development (Ambros and Horvitz 1984; Ambros 1989; Ambros and Moss 1994). Mutations in each of these genes cause specific cells to adopt roles that their lineage normally plays earlier or later (Ambros and Horvitz 1984). Some of these mutations correspond to differences observed among nematode species (Ambros 1988). Yet, in each species these genes interact in a coordinated way to specify the timing of the switch from larval to adult molts (Ambros 1989; Ambros and Moss 1994). Homeotic genes may have played a similar role in the evolution of insects (Lerclerc and Regier, 1990) and vertebrates (Duboule 1994).

Simple changes in the endocrine growth control can produce heterochronic effects at the whole-organism level (Shea 1992a). Merimee et al. (1987) showed that short stature in African pygmies mainly results from the reduction of the adolescent growth spurt and related this to the very low levels of insulin-like growth factor 1 (IGF I) during puberty in pygmies relative to other ethnic groups (see also Shea, 1989; Shea et al. 1990; McKinney and McNamara, 1991). The endocrine control of metamorphosis in insects is fairly well understood (Nijhout 1994a), and can provide the basis for complex polymorphisms (e.g., Wheeler 1990, 1991).

Plants provide an interesting system for the study of ontogeny because of their modular architecture, in which elements such as leaves or flowers are repeated in a series along a shoot axis (Guerrant 1988). In general the position on the shoot corresponds to the time of initiation of a structure, and the differences among leaves or flowers iterated along an axis (heteroblasty) can thus be interpreted as a temporal record of whole-plant ontogeny. In this context, heterochrony has been invoked as a possible origin of cleistogamous flowers (Lord and Hill 1987; Gallardo et al. 1993) and to explain differences in leaves be-

tween wild species and cultivars (Jones 1992, 1993; but see McLellan 1993). Wiltshire et al. (1994) reported on several mutants of the garden pea that have clear heterochronic effects, suggesting that at least some heterochronic changes may originate from fairly simple genetic and developmental changes.

Some of the preceding examples demonstrated that heterochronic changes can have a simple genetic basis. Even in the examples where effects of individual genes are known, however, it may be more appropriate to view genes as suppliers of materials used in a complex network of developmental interactions, rather than as the direct cause or controlling agents of development (Nijhout, 1990b; Alberch, 1991). Change of any single part in this network may produce a cascade of effects, but development relies on the interplay of all the parts. These intricate interactions among developmental processes that produce heterochrony make it difficult to identify and understand the mechanisms underlying heterochrony.

As an alternative, one can resort to statistical analysis of genetic and developmental phenomena using the methods of quantitative genetics. A number of quantitative genetic models for the evolution of ontogenies have been proposed (Atchley, 1987; Slatkin, 1987; Atchley and Hall, 1991; Cowley and Atchley, 1992; Atchley et al. 1994). These models are often very complex themselves, and empirical studies have therefore focused either on the covariation among traits in adults, which does not relate directly to heterochrony, or they have analyzed genetic variation and constraints in the growth functions of one or more traits (e.g., Cheverud et al. 1983; Atchley 1984; Leamy and Cheverud 1984; Riska et al. 1984; Lynch 1988; Kirkpatrick and Lofsvold 1989, 1992).

Analyses of phenotypic or genetic covariation among size measurements of size at various ontogenetic stages can indicate the potential for evolutionary change in ontogenies. Such studies have demonstrated that there is a high degree of covariation among stages, with little flexibility for independent variation in different stages (Cheverud et al. 1983; Riska et al. 1984; Kirkpatrick and Lofsvold 1992; Björklund 1993; Klingenberg 1996b). The patterns of covariation in age-specific size show clear commonalities among these studies of mammals, birds, and insects; in contrast, corresponding analyses of growth increments show no such coordinated variation in insects and rodents (Klingenberg 1996b). This suggests that the strong patterns found in the analyses of age-specific size data result from the part-whole relationships inherent in these cumulative data, but do not reflect specific properties of growth processes in these organisms.

Measures of Time

Timing and rates of development are the central concepts of heterochrony. While the importance of time is clear intuitively, and is generally acknowledged in the literature, it is far less evident how the age of an organism should be measured. There is no agreement on whether such a metric should reflect developmental or physiological processes within the organism (intrinsic time) or, conversely, whether it should be independent of them (extrinsic time). Moreover, a variety of different measures of intrinsic time have been proposed in the literature. The formalisms for analyzing heterochrony all require a measure of time, but differ in the ways they incorporate it.

Comparisons of ontogenies with the clock model require identifying two homologous ontogenetic stages in both ancestral and descendant species (for diagnosing the type of heterochronic change, one stage plus information about ontogenetic polarity are sufficient). The first stage gives initial size, shape, and age to calibrate the scales of the clock (the "zero" point), and the second is the standard stage at which ancestor and descendant are compared. Size, shape, and age are measured at these two stages and possibly at intervening stages, and heterochrony is assessed by displaying the descendant ontogeny on the scales calibrated for the ancestor. Provided the homologous stage for the comparison can be identified reliably for both ancestor and descendant, the conclusions drawn from the clock model are relatively robust against differences in the choice of a metric for age, because the different types of heterochrony are identified by qualitative comparisons of age, size, and shape (younger–older, smaller–larger, paedomorphic–peramorphic; see Fig. 3-1).

The formalism based on growth functions, however, explicitly refers to developmental rates and to the time of onset or termination of growth processes, and therefore requires quantification of these parameters using a metric of time. In addition, one stage must be identified as homologous among the species being compared, at initiation of development (at age zero). This is usually not too difficult, taking into account the context of a particular study. For instance, the time of fertilization can be defined as this reference stage for studies dealing with embryonic development, hatching or birth for postembryonic growth. The choice of a metric for time, however, is more problematic because different measures of time often are related in a highly nonlinear way. Alberch et al. (1979:301) specified explicitly that age and time advance at the same rate; this means that extrinsic time is to be used as the framework of analysis. Nonetheless, the use of extrinsic time for comparative purposes has been criticized because developmental rates depend on environmental factors such as temperature and on the size of the organism itself (e.g., J. O. Reiss 1989). Some authors even proposed that size may be a more appropriate measure of ontogenetic "age" (e.g., Strauss 1987, 1990).

Intrinsic and extrinsic time

The most fundamental division among the several ways in which the passing of time can be measured is that between intrinsic and extrinsic time. *Intrinsic time* is also variously called physiological time or developmental time, depending on the particular context. It measures time by processes within the organism, for example by the occurrence of discrete developmental events. Therefore, progression of intrinsic time is sensitive to the size and other properties of the organism itself and to environmental factors such as temperature and nutrition—time passes at different rates for different organisms. In contrast, the advance of *extrinsic time*, also called clock time or astronomic time, is independent of an organism's condition or environment, but rates of biological processes measured on this time scale may fluctuate within organisms according to conditions.

In principle, at least, the choice of intrinsic or extrinsic measures of time is a matter of convenience (Hall and Miyake 1995), because time can be converted from one measure into another, provided sufficient information is available. In practice, however, this is hardly ever the case, because the relationships be-

tween different measures of time are often nonlinear and dependent on multiple physiological and ecological factors.

The scales of intrinsic and extrinsic time can be remarkably incongruent. Despite controlled laboratory conditions, the durations of the five larval instars varied by a factor of about two among individuals of a water strider species (Klingenberg 1996b). Hall and Miyake (1995:11–14) showed that the relationship between morphologically defined stages of mouse embryos changes markedly during development, with ample variation in early stages and later catch-up phenomena.

The simplest expressions of intrinsic time are stages defined by discrete events, such as hatching or birth, molts, or maturity; more elaborate schemes are used in tables of normal development (Hall and Miyake 1995; Starck 1993) and more recently in studies of the order of gene expression (Duboule 1994). These only record the sequence of events, but cannot quantify the intervals between them. Modular organisms, such as vascular plants, offer the opportunity to use a very different measure of intrinsic "time" without even measuring time per se (Ritterbusch 1990). The counts of modules, for example the nodes with leaves along a shoot of a plant, lay down a record of growth processes that can be read from the structure of the mature organism (e.g., Jones 1992; Wiltshire et al. 1994). Richardson (1995) used somite counts in a similar way to compare stages of embryonic development in vertebrates.

If measurements of the intervals between successive events are of interest, a unit for intrinsic time must be chosen, which often involves standardization for environmental conditions and size. The simplest metric expresses these intervals as a percentage of a certain period, such as embryonic development from oviposition to hatching. Physiological time is most commonly used to control for the influence of environmental factors in intraspecific studies, especially to correct for temperature variation in ectothermic animals (e.g., Taylor 1981; Pruess 1983; Sinervo and Doyle 1990). Body size is another factor that strongly influences the duration of most phases of the life cycle (e.g., Calder 1984). After correction for these influences, interspecific comparisons examine if development of a particular organism is faster or slower than would be expected under the given conditions from a comparison with the "average" of many related species. Some measures of intrinsic time have been developed specifically for comparisons of the development of different species. They measure time in units of cell cycle durations (Dettlaff 1986; Dettlaff et al. 1987) or as the accumulation of metabolic activity per unit of body mass (J. O. Reiss 1989).

Adjustments for physiological time should be made carefully, because they may eliminate variation that is of interest in studies of heterochrony. Imagine an organism that grows larger than its ancestor by growing for a longer period as measured in units of extrinsic time. On this time scale, this heterochronic change would be clearly diagnosed as hypermorphosis. If developmental stages linked to maturity are used to determine units of intrinsic time, the descendant has the same intrinsic age at cessation of growth by definition. On this time scale, consequently, it must have a higher growth rate to reach its larger size, and the heterochronic change would be diagnosed as acceleration. Klingenberg and Spence (1993: Fig 9) showed that extrinsic time and discrete developmental events (molts) lead to different conclusions about heterochrony in water striders. Similar problems apply to corrections for temperature effects

and the choice of the experimental temperature in laboratory studies if species differ in their temperature optima. Blackstone (1987b) argued for the use of extrinsic time because it provides an unambiguous standard in comparative studies (see also Blackstone and Yund 1989; McKinney and McNamara 1991; for the opposite viewpoint, see Strauss 1987).

Incremental growth marks in hard tissues provide a connection between intrinsic and extrinsic time. Changes in growth rates can leave marks in shells, bone or teeth, and these fluctuations are often cyclical and synchronized with seasonal changes or other periodical variation in the environment. Such marks, laid down at regular intervals of extrinsic time, can be related to the development of the organism and its intrinsic time scale. Hard parts are increasingly being used for aging of fossil and recent animals (Jones 1988; Castanet et al. 1993). For instance, studies of growth markings in teeth have been instrumental in reconstructing the life histories of fossil hominids (Beynon and Dean 1988; Smith and Tompkins 1995).

Size as time: "Allometric Heterochrony"

In many studies of heterochrony, especially in fossil organisms, age data are not available. Because size increases monotonically with age during growth of most organisms, it seems straightforward in this situation to substitute size as a measure of intrinsic time. Strauss (1987:73) even proposed that overall body size is an estimate of biological age more directly tied to growth than chronological time. Following similar logic, McKinney (1986) formalized a framework of "allometric heterochrony" using bivariate allometric plots to infer heterochronic processes (see also Gould 1982:336; McKinney 1988; McKinney and McNamara 1991). The basic assumption is that size increases according to a growth schedule identical in ancestral and descendant forms being compared; moreover, in this scheme a trait measurement (e.g., the size of a particular organ) is used as the measure for "development" in the modified version of the framework of Alberch et al. (1979; see Fig. 3-3). Under these conditions, the formalism of allometric heterochrony predicts that acceleration will lead to a higher slope, and neoteny to a reduced slope. Progenesis will produce an allometric trajectory of the descendant that follows the ancestral one, but is truncated, whereas hypermorphosis will yield an extension of the ancestral trajectory; both of these allometric result are cases of ontogenetic scaling (Gould 1975). Predisplacement means that the trait starts growing earlier in the descendant than in the ancestor, while size still follows the ancestral schedule; for any given size value, the trait is therefore larger in the descendant than in the ancestor, and the allometric plot of the descendant has thus a higher *y*-intercept than that of the ancestor. Conversely, postdisplacement leads to a lower descendant *y*-intercept under these conditions. Numerous authors have used this approach to interpret allometric plots in terms of heterochrony (e.g., Alberch and Alberch 1981; Gould 1982:336; McKinney 1986; McNamara 1988; Lessa and Patton 1989; Winterbottom 1990; Boughton et al. 1991; Allmon 1994; Vrba 1994; Vrba et al. 1994; Wei 1994).

McKinney (1986, 1988) and McKinney and McNamara (1991) discussed a number of possible problems with equating size and age and stated a number of caveats for users of allometric heterochrony. Still, they maintained that the diagnoses are correct if the growth dynamics of size are the same in ancestor and descendant (McKinney 1988:21f.; McKinney and McNamara 1991:37). A

number of other authors have presented more far-reaching criticism of allometric heterochrony. These critiques include empirical investigations demonstrating failure of the method in particular cases as well as theoretical arguments addressing problems with the underlying logic.

Some empirical studies have demonstrated evolutionary changes in the growth dynamics of overall size, which can generate incongruence between the findings of age- and size-based methods for identifying heterochrony (e.g., Emerson 1986; Blackstone and Yund 1989; Klingenberg and Spence 1993; Leigh and Shea 1996). The monotonic increase of size with age of each individual does not imply that a corresponding relation also exists among individuals or even across different taxa. For example, analyses in water striders showed that development time and adult size are negatively correlated within populations and uncorrelated among species (Klingenberg and Spence 1993, 1996; Klingenberg 1996b). This lack of correlation is due to independent variation in the growth rates of size and shape features—and this failure of the bacterium-to-whale regressions at a lower phylogenetic scale in turn constitutes the essence of evolution by heterochrony (see discussion by Blackstone and Yund 1989; McKinney 1988; McKinney and McNamara 1991:35–40; Klingenberg and Spence 1993).

The allometric patterns expected are not unique to particular heterochronic processes. For instance, not only progenesis and hypermorphosis, but also changes in growth rates or onset time can produce ontogenetic scaling, provided they affect all measured organs simultaneously. Such concurrent changes have been amply documented in studies of primates (Shea 1983a, 1988, 1989) and dog breeds (Wayne 1986a, 1986b). Evolutionary changes by ontogenetic scaling may be widespread, because changes in hormonal growth control provide a simple physiological basis for these coordinated alterations in growth rates (Shea 1992a).

The empirical evidence of heterochronic changes in both size and shape is the starting point for theoretical arguments showing that many different changes can lead to the same pattern. For example, Blackstone and Yund (1989:8) listed five different ways that can lead to an increase in the slope of an allometric plot. Moreover, Klingenberg and Spence (1993) demonstrated with a simple graphical model (Fig. 3-6) that there is no consistent set of conditions that leads to the patterns of allometric heterochrony. This model simulates simple heterochronic changes (i.e., in only one of the parameters in the model of Alberch et al. 1979; see Fig. 3-3) for a trait, overall body size, or both. Neoteny and acceleration are only correctly identified if they affect the trait, but not overall size. In contrast, the allometric plots for predisplacement and postdisplacement only conform to the expected patterns if the first growth period, at the time of onset, is ignored. Finally, progenesis and hypermorphosis only result if both the trait and size change simultaneously, but any heterochronic change in both variables produces one of these two patterns. In sum, the expected allometric patterns occur only under specific conditions that depend on the heterochronic changes themselves, and knowledge of heterochronic processes is therefore necessary to diagnose them correctly.

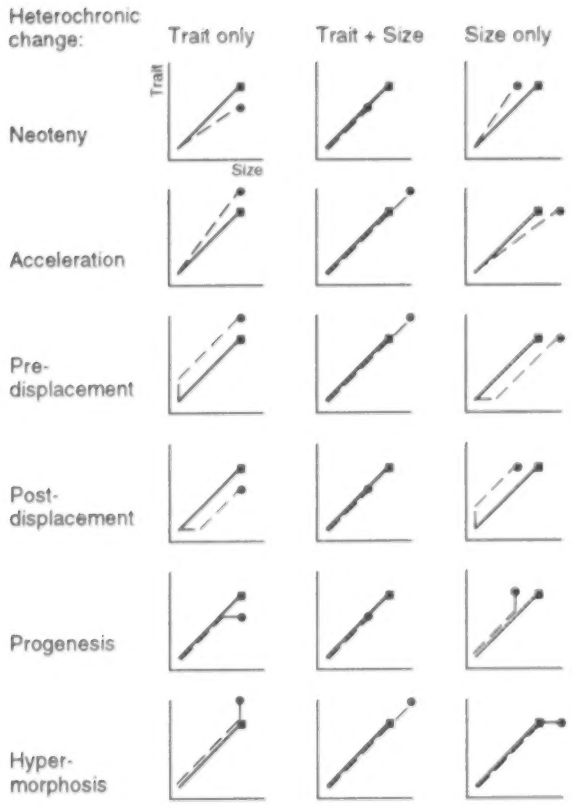


Fig. 3-6. Heterochronic changes and their effects on allometric plots of a trait measurement versus overall size. The allometric patterns expected under McKinney's (1986) "allometric heterochrony" are only found if neoteny or acceleration only affect the trait alone, if progenesis or hypermorphosis affect both size and the trait simultaneously, or if predisplacement and postdisplacement affect the trait alone and the horizontal or vertical parts of the growth trajectories have not been observed (i.e., a lack of data during the part of the ontogeny when the change is actually occurring). For simplicity, I assume that ancestral onset and offset times are the same for size and the trait, and that all heterochronic changes have effects of the same relative magnitude on both variables (see Fig. 3-3; note that growth functions need not be linear, but throughout the growth period, the specific growth rates of the two variables are always proportional, or one of them is zero during some part of ontogeny if there are changes of the onset or offset parameters). These assumptions can be relaxed to some extent without affecting the results substantially. (From Klingenberg and Spence 1993, with permission of the Society for the Study of Evolution.)

In a recent series of papers, Godfrey and Sutherland (1995a, 1995b, 1996) presented further criticism of allometric heterochrony. They emphasized the contradictions between interpretations of heterochrony derived from allometric plots of a trait y versus a size variable x and those derived from Gould's clock

model if y/x is taken as the shape variable (Godfrey and Sutherland 1995b, 1996). In this case, the patterns expected under McKinney's allometric heterochrony hold only if y grows with positive allometry relative to x, but with negative allometry, the direction of changes in allometric plots reverses itself (see also Shea 1989:73). The use of logistic growth functions and specific definitions of size and shape in their numerical simulations limits the generality of some of Godfrey and Sutherland's results, because changes in single parameters of the logistic equation can produce complex changes in the growth curves (cf. Fig. 3-5). In a similar model of sigmoid growth in two traits, using the Gompertz function, simple relative shifts of the growth curves along the time axis produced a change in the slope of the allometric plots, a pattern that would be misdiagnosed as neoteny or acceleration by the framework of allometric heterochrony (Laird et al. 1968; Barton and Laird 1969).

McKinney and McNamara (1991:41) recognized that the allometric diagnoses will not correspond to the true heterochronic process under some circumstances, and they recommended to use allometric heterochrony as a concept distinct from age-based heterochrony, by adding the qualifier "allometric" each time one of the heterochronic terms is used. The theoretical arguments presented above show that the diagnoses of allometric heterochrony can be incompatible with the true heterochronic processes even if the caveats of McKinney (1988) and McKinney and McNamara (1991) are taken into account (see Fig. 3-6), and the empirical studies suggest that this commonly is so. Therefore, it seems highly questionable to call an increase in slope "allometric acceleration" when all this implies is an increase in slope! Allometric analyses themselves provide a rich opportunity to investigate ontogenies and their evolutionary modification, and their terminology should be kept separate from the distinct but complementary concept of heterochrony.

ALLOMETRY

The concept of allometry, like heterochrony, has several different meanings and multiple methodological approaches are available for analyses. All have in common that allometry deals with covariation of traits associated with variation in overall size of organisms. The traits can be the size of parts, shape, or physiological, ecological, and behavioral characteristics, but the range of traits considered differs among the various concepts of allometry. In this review, I concentrate on morphological features.

Allometry fundamentally differs from heterochrony in that it does *not* explicitly include the dimension of time in the analysis (Huxley 1932; Reeve and Huxley 1945; Laird 1965; Cock 1966; Gould 1966; Shea 1985; Blackstone 1987a, 1987b; Klingenberg 1996a). The domain of allometry is purely morphological, and concerns measures of size and shape. The temporal dynamics of growth enter the analysis of ontogenetic allometry indirectly, because the growth curves determine the value of each morphological measurement at every age (Fig. 3-1). German and Meyers (1989a, 1989b) did not make this distinction between growth functions and allometry when they discussed the choice between size and age as independent variables for allometry; their "growth allometry" of weight versus age (1989b: Table 1) is a growth curve following a power function, but it is clearly not allometry in the sense of this review and other authors.

Time can be included as a dependent variable in allometric comparisons at the interspecific level (e.g., Calder 1984; M. J. Reiss 1989). In these kinds of studies, however, size is always the independent variable, and is used to account for variation in temporal parameters characteristic for each species (e.g., total lifespan, age at maturity).

Size, Shape, and Allometry: Two Schools of Thought

The concept of allometry has developed over time (Blackstone 1987a, 1987b), and there are currently two major conceptual frameworks of allometry, which have different implications for the connection between heterochrony and allometry. These two approaches differ mainly in the way they define and analyze organismal size and shape, and therefore reflect the spectrum of methods in morphometrics (e.g., Rohlf and Bookstein 1990). They are not as well defined as the various formalisms for analyzing heterochrony, and there are some methodological approaches that bridge the gap between them, which is probably the reason why they have not been recognized and compared (but see Bookstein 1989; Klingenberg 1996a). Empirical assessments of allometric methods usually have adopted one of these frameworks as the standard for comparison; it is thus hardly surprising that such comparisons usually criticize one approach because for its failure to produce results deemed "correct" under the alternative framework (e.g., Albrecht et al. 1993; Jungers et al. 1995).

The first approach to the study of allometry, which I call the Huxley–Jolicoeur school, is in the tradition of Huxley's work, which was based initially on a model of growth dynamics. In this framework, allometry is the pattern of covariation among parts, and organismal shape is defined informally as the relative sizes of parts. The second line of thought, which I call the Gould–Mosimann school, rigorously defines shape by geometric similarity; allometry is any association between size and shape, but does not refer to a biological process explicitly. (I have named the two approaches after authors who made seminal contributions to the conceptual foundations or analytical methods currently in use. I do not imply, however, that these researchers exclusively used one or the other of these approaches.)

The Huxley–Jolicoeur School

The most widely used expression for allometry is the equation of *simple allometry* proposed by Huxley (1924, 1932): $y = bx^k$ or equivalently $\log y = \log b + k \log y$. Huxley originally introduced this equation to study relative growth, and it is still routinely used in this context. If $k > 1$, y is called *positively allometric* with respect to x, and the ratio y/x will increase through the growth period. Conversely, there is *negative allometry* if $k < 1$, and the value of y relative to x will decrease with growth. The coefficient k, which provides information about ontogenetic change in the relative magnitude of y versus x, can therefore be used to establish ontogenetic polarity for analyses of heterochrony. If $k = 1$, the traits are *isometric* and only the absolute sizes of x and y change during growth, because the ratio between x and y is constant (i.e., $y = bx$), and there consequently is no ontogenetic polarity.

When Huxley (1924, 1932) introduced his equation of simple allometry, he specifically referred to a process of multiplicative growth. The allometric coefficient k is determined by the ratio of the specific growth rates of two traits:

$$k = \frac{\frac{dy}{dt}\Big/y}{\frac{dx}{dt}\Big/x}$$

(Huxley 1932:6; Reeve and Huxley 1945; Laird 1965; Lande 1985; Shea 1985). As long as the specific growth rates of both traits have a constant ratio, the resulting allometric plot will be linear on a log-log scale.

A range of conditions can lead to constant ratios between the specific growth rates of a pair of traits, and therefore simple allometry can be the consequence of a number of biological phenomena (see also German and Meyers 1989a). Katz (1980) used a simplified model of cell proliferation to derive the allometry equation. Laird and coworkers used a phenomenological model of growth based on the Gompertz function, which assumes an exponential decay of initially high specific growth rates (Laird 1965; Laird et al. 1968; Barton and Laird 1969). This model generates simple allometry as long as the coefficient of decay of the growth rates is the same for both variables. Blackstone (1986, 1987c) directly analyzed specific growth rates and their ratios in hermit crabs.

Extending the concept of simple allometry to multiple measurements is fairly straightforward. For many data sets, examination of growth in several traits shows that all pairwise allometric plots are linear or nearly so; assuming linearity of all pairwise allometric plots among multiple measurements provides a direct multivariate extension of simple allometry (Klingenberg 1996a). Each of these plots can be considered as a projection from the space spanned by all the measured variables onto the plane defined by two variables. For three variables x, y, and z, it is easy to see that the growth trajectory in the three-dimensional space must be a straight line if all three projections onto a plane (x vs. y, x vs. z, and y vs. z) produce linear allometric plots. A similar argument holds for more than three dimensions (Sprent 1972). Under this concept of multivariate simple allometry, growth trajectories are straight lines in the space of log-transformed measurements.

The statistical task for analyzing allometry is therefore to find a line of best fit to the scatter of data points in this multidimensional space. Jolicoeur (1963) proposed to use the first principal component of the covariance matrix of log-transformed measurements as a multivariate generalization of allometry. Klingenberg (1996a) reviewed biological and statistical aspects of this approach and summarized recent extensions for comparisons among groups.

There can be curvatures of growth trajectories, indicating that the ratios of specific growth rates vary through ontogeny (e.g., Huxley 1932; Bookstein 1991: Figs. 4.2.2–4.2.4; Klingenberg and Zimmermann 1992; Klingenberg and Spence 1993). Nonlinearities can occur in cases where the growth rates of certain structures change drastically relative to the rest of the body (Cuzin-Roudy and Laval 1975; Boag 1984; Cane 1993). In animals, these bends in growth trajectories are often associated with events such as the transition from larva to adult (e.g., Strauss and Fuiman 1985). In contrast, nonlinear allometries seem

to be fairly widespread in plant development (Kampny et al. 1993, 1994; McLellan 1993).

The concept of "size" is an important component of this approach to allometry. Unless a measure of size is specified a priori in a bivariate analysis, it must be derived from the variables in the study. Because the position along the growth trajectory is an intuitive measure of overall size, such "size scores" have been used traditionally in multivariate studies (e.g., Jolicoeur and Mosimann 1960; Bookstein et al. 1985; Bookstein 1989; Klingenberg 1996a). A "size" variable occurs in almost every ontogenetic data set, and often accounts for most of the variation (99% is not exceptional; see Klingenberg 1996a). Moreover, these size scores are highly correlated with other possible size measures, such as individual measurements (e.g., body weight, "standard length" for fish) or the arithmetic or geometric mean of all variables.

In contrast, the notion of shape is only of peripheral importance in the Huxley–Jolicoeur approach to the study of allometry: the analysis is concerned with the covariation among sizes of parts, but not directly with their proportions, which constitute shape in its vernacular sense. There is a link to shape, however, because if growth is allometric (with any bivariate $k \neq 1$), at least some proportions among variables will change during growth.

In general, shape is fairly difficult to deal with in this framework. Many workers have considered any variation as "shape" that is not "size", and thus defined "shape" as all variation in directions orthogonal to the "size" axis (e.g., Jolicoeur and Mosimann 1960; see discussions by Bookstein 1989; Klingenberg 1996a). This concept of "shape," however, is far removed from the idea of shape related to proportions and geometric similarity. If there is allometric variation, proportions among variables will change along the allometric axis, which therefore contains not only variation of size but also of shape. Therefore, this method separates size and size-related variation in proportions along the allometric axis (i.e., not "size alone"), from other, size-independent "shape" variation (see Flessa and Bray 1977). In many situations this is desirable, for instance, if a researcher attempts to correct for growth variation when comparing two groups (see Burnaby 1966; Klingenberg 1996a); however, the terms "size" and "shape" should be used only with caution in discussions of results from such analyses.

For studies of heterochrony, this approach to allometry is most compatible with the modified version of the formalism of Alberch et al. (1979), in which size and shape are not necessarily separated. In these studies, the position along the growth axis also provides an ideal way to establish ontogenetic polarity in size and the principal ontogenetic shape changes (e.g., Creighton and Strauss 1986; Klingenberg and Spence 1993).

The Gould–Mosimann School

In his review of the subject, Gould (1966) expanded the definition of allometry to mean "the study of size and its consequences" (p 587). Moreover, he explicitly separated allometry from any specific mathematical relationship among variables, and thereby denied any special status of the simple allometry equation besides the fact that it often fits empirical data well. For morphology, therefore, allometry merely implies that there is some shape change associated with increase in size; conversely, the absence of size-related shape variation is

isometry (in perfect agreement with the Huxley–Jolicoueur approach). Mosimann (1970) proposed a formal statistical definition of this concept of allometry: allometry is an association between shape and size, whereas isometry is their stochastic independence.

In addition to this revision of the concept of allometry, Mosimann (1970) also provided a mathematical framework for the analysis of size and shape based on geometric similarity. The size and shape of an object are characterized by a vector of measurements. Two objects have the same shape if multiplication of the measurement vector of one object with a positive constant can transform it into the vector of the other object. Whereas shape is inherently multivariate, size is a scalar; Mosimann defined a size variable as any function of the measurement vector that scales linearly (i.e., multiplying every measurement with a constant yields a value of the size variable that is multiplied by the same factor). Photographic reduction or enlargement of an image are analogs to this concept of shape equality; the magnification indicates the change of a size measure.

In bivariate allometric plots on log-log scales, the locus of geometrically similar organisms is a straight line with a slope of 1. Similarly, in a multivariate context, it is a straight line at equal angles to all coordinate axes in the space of log-transformed traits (i.e., with a coefficient vector that is a scalar multiple of $[1, 1, 1, \ldots, 1]'$). It is therefore straightforward to use the position along these lines to measure size. In the context of Mosimann's (1970) theory, this means that one chooses the geometric mean of all variables as the size variable. All variation in directions perpendicular to this isometric axis involves differences in geometric shape.

Numerous authors have used this approach to analyze variation in geometric shape. The method has been proposed several times independently under a variety of names; the different implementations differ in the way calculations are done, but are equivalent except possibly for minor details such as overall centering. The first implementation of the method projects the data points onto a subspace orthogonal to the isometric vector (Burnaby 1966; Rohlf and Bookstein 1987; Somers 1989:171) whereas the alternative approach doubly centers the matrix of log-transformed data to have means of zero for both rows and columns (Mosimann and James 1979; Darroch and Mosimann 1985; Kazmierczak 1985; Berge and Kazmierczak 1986; Somers 1989:171; Berge 1991; Yoccoz 1993). A similar method uses residuals from regression of each measurement on the isometric size variable to eliminate size variation (Healy and Tanner 1981; Wilson and Loesch 1989).

Euclidean distance matrix analysis (EDMA) has been proposed recently as a new method to compare objects in two or three dimensions (Lele and Richtsmeier 1991; Richtsmeier and Lele 1993; Richtsmeier et al. 1993). The technique analyzes the matrix of all pairwise distances between a set of morphological landmarks. EDMA compares pairs of forms or sample averages by calculating ratios of corresponding distances; by comparison of juveniles to adults it can thus test if growth is isometric. Moreover, EDMA can compare the patterns of juvenile-adult growth ratios among pairs of species. Nevertheless, the flexibility of the technique is somewhat limited, and analyses involve large numbers of variables even with moderate numbers of landmarks (e.g., 10 landmarks result in 45 pairwise distances). Therefore, it is uncertain how this technique can be applied to heterochrony. Although Richtsmeier and Lele

(1993) discussed this topic, the only worked example (their Fig 12) used only three landmarks (of the 10 or more in EDMA examples), so that all pairwise distances between them could be plotted in the space of log-transformed measurements; thus the example followed the methodology outlined in the preceding section.

Another group of methods, often referred to as "geometric morphometrics," defines shape by the configuration of a set of morphological landmarks, rather than by the proportions among distances between landmarks (Rohlf 1990; Rohlf and Bookstein 1990; Rohlf and Slice 1990; Bookstein 1991, 1993). This approach provides a flexible array of methods for superposition and comparison of shapes, and for the multivariate analysis of shape change. The drawback of these methods for analyzing ontogenies is the complexity of growth trajectories in the resulting shape spaces and their high dimensionality. Allometric trajectories are often markedly nonlinear (e.g., Bookstein 1991: Figs. 7.6.4–7.6.7; Zelditch et al. 1992, 1993; Walker 1993; Zelditch and Fink 1995), and they are therefore difficult to analyze, for example, with regard to extension of growth by ontogenetic scaling.

Because the techniques described in this section use an explicitly geometrical definition of shape, the resulting shape variables can be used as measures of shape for analyses of heterochrony. As shape is inherently multidimensional, however, different shape variables may show different ontogenetic trends and the results of comparisons will vary correspondingly. Summary vectors to establish ontogenetic polarity may be computed with multivariate analyses of shape data, or by regression of size or age data on the shape variables (using multivariate and nonlinear regression techniques).

Evolution of Ontogenetic trajectories

Ontogeny can be described by the allometric trajectory of an organism and the rate at which it proceeds along the trajectory. Heterochrony is evolutionary change in these trajectories and rates. Both allometric analyses and timing are therefore crucial for the understanding of heterochrony.

Allometric trajectories can be extended or truncated, they can change their direction, or they can be shifted sideways (Fig. 3-7). Comparisons of trajectories can provide information about the dissociation of the growth dynamics among measurements. Changes in direction of allometric trajectories indicate such dissociation during the ontogenetic phase being studied, lateral transposition indicates dissociation in an earlier period, whereas conserved trajectories indicate the maintenance of ancestral associations among traits. This information is relative, and joint changes in the rates or duration of growth that affect all traits will not be discovered with allometric analyses alone.

Extension and truncation of ancestral ontogenies have been discussed under the heading of ontogenetic scaling (e.g., Gould 1975; Shea 1988, 1992b; Ravosa et al. 1993). Ontogenetic scaling, when it is found without other changes to growth trajectories, can establish paedomorphosis or peramorphosis directly and unambiguously, but age is essential to identify the heterochronic process responsible. Some authors have associated ontogenetic scaling with progenesis and hypermorphosis, that is, changes in the age at termination of growth (McKinney 1986; McNamara 1988), but a number of studies have shown that other kinds of coordinated changes in growth dynamics of multiple

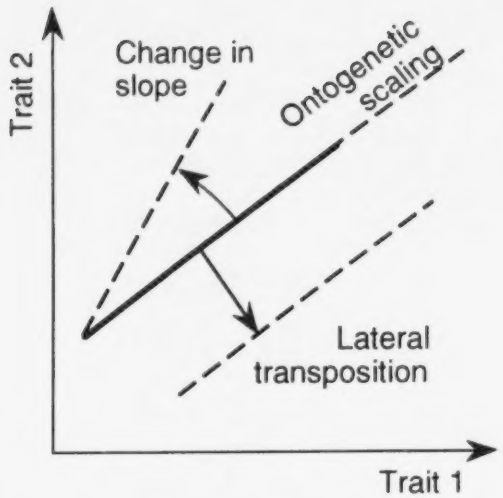


Fig. 3-7. Evolutionary changes to allometric trajectories. Ontogenetic scaling is the extension or truncation of the ancestral trajectory. Changes in direction (slope in this bivariate graph) and lateral transposition (i.e., a parallel shift of the entire trajectory) involve dissociation of the growth schedules of the traits. The ancestral allometric trajectory is drawn as a solid line, and dashed lines represent several descendant allometries.

measurements can produce the same allometric pattern (Shea 1983a, 1988, 1989, 1992b; Wayne 1986a, 1986b; Shea et al. 1990; Ravosa et al. 1993).

Changes in directions of growth trajectories indicate that the between-trait ratios of specific growth rates differ from ancestor to descendant (for at least one pair of variables), that is, growth dynamics have become dissociated among traits. Changes in the directions of allometric trajectories can lead to paedomorphosis and peramorphosis, but this ontogenetic polarity of change depends on ancestral allometry, as pointed out by Shea (1989:73) and Godfrey and Sutherland (1995b, 1996) for bivariate allometry (see the critique of allometric heterochrony, above). I propose a simple thought experiment to explain this fact for bivariate and multivariate allometry, which also illustrates the connections between the Huxley–Jolicoeur approach and geometric concepts of shape. Imagine a descendant maximally paedomorphic according to geometric shape: instead of following the ancestral trajectory towards more adult-like proportions, this descendant would retain the starting shape and grow isometrically without change of geometric shape. Any change in direction of the allometric trajectory that brings it closer to the isometric vector $(1, 1, \ldots, 1)'$ would therefore produce paedomorphosis, and peramorphosis would result from increases in the angle between the trajectory and the isometry vector. This reasoning explains the simulation results of Godfrey and Sutherland (1995b, 1996).

The angles between ancestral and descendant trajectories and between each of them and the isometry vector therefore provide a rough measure for ontoge-

netic polarity (see, e.g., Klingenberg 1996a). Note, however, that the interpretation of ontogenetic polarity in a multivariate context is difficult if the descendant trajectory is not within (or at least near to) the plane defined by the ancestral trajectory and the isometric vector through the starting form. Moreover, these angles are difficult to interpret if ancestral and descendant trajectories differ more from each other than from the isometry vector, because ontogenetic polarity in the ancestor and descendant may be opposite (e.g., a switch from positive to negative bivariate allometry); as in the special case of bivariate allometry, this polarity is undefined if the ancestral trajectory is isometric.

Lateral transpositions indicate that the starting forms differ between ancestor and descendant. This means simply that alterations in development occur in early developmental stages not included in an analysis. Nevertheless, as ancestral and descendant growth trajectories are parallel in the space of log-transformed measurements, the dynamics of growth in all traits must preserve the ancestral associations. Larval growth of water striders provides a striking example of this conservation of allometric trajectories: although there are extensive lateral transpositions, even the subtle deviations from simple allometry are the same from species to species (Klingenberg and Spence 1993). In a multivariate context, the technique of Burnaby (1966) allows one to separate variation into components parallel and orthogonal to the growth trajectory, and is therefore ideal to study lateral transposition and ontogenetic scaling even if they occur together; a detailed discussion and applications are provided by Klingenberg and Spence (1993) and Klingenberg (1996a).

Empirical studies considering multiple traits and carried out with sample sizes large enough to provide adequate statistical power often have found a combination of all these processes (e.g., Shea 1983, 1985, 1989; Wayne 1986a, 1986b; Klingenberg and Spence 1993; Klingenberg and Ekau 1996).

Static, Ontogenetic, and Evolutionary Allometry

Allometric analyses can address variation at several levels, which the different biological origins of variation and covariation among traits. These levels of allometry can be classified in several ways (Cock 1966; Gould 1966, 1975). I follow Cock (1966) in distinguishing static, ontogenetic, and evolutionary allometry (see also Cheverud 1982; Klingenberg and Zimmermann 1992; Klingenberg 1996a).

The preceding sections focused on ontogenetic allometry, for which growth is the source of morphological variation. For most ontogenetic data sets, the model of simple allometry fits well, and the model of a linear growth trajectory in the space of log-transformed characters can therefore be used to compare the different levels of allometry.

Static allometry reflects trait covariation among individuals at a particular ontogenetic stage and within a single population (Cock 1966; Klingenberg and Zimmermann 1992; Klingenberg 1996a). It is "static" in that it represents a snapshot of individual variability that eliminates the influence of both ontogenetic and evolutionary dynamics. Gould (1966, 1975) used the term "intraspecific" for this level of allometry (note that Gould used the term "static" for interspecific comparisons, which I treat as evolutionary allometry). The ontogenetic stage at which static variation is analyzed is most often the adult, but studies in other stages are equally informative; comparisons between

stages of static variation can provide insight into the ontogenetic changes of individual variability (e.g., Cuzin-Roudy 1975; Zelditch 1988; Zelditch and Carmichael 1989; Klingenberg and Zimmermann 1992; Klingenberg et al. 1996).

A number of empirical studies have compared static and ontogenetic allometry. Generally, correspondence between these allometric patterns is at least fairly close (Cheverud 1982; Leamy and Bradley 1982; Klingenberg and Zimmermann 1992; Klingenberg 1996b), but one study found considerable differences related to traits with early or late growth maxima (bill versus wings in Darwin's finches; Boag 1984). These studies suggest that much of the variation among individuals stems from a variable extent of growth along relatively constant allometric trajectories (see discussions by Cock 1966; Cheverud 1982; Klingenberg and Zimmermann 1992; Klingenberg 1996b). Patterns of static allometry can provide evidence regarding the underlying developmental processes, especially if this is supported by physiological and genetic experimentation (e.g., Wheeler 1990, 1991; Emlen 1994). Riska (1986) presented theoretical models of developmental processes that generate various patterns of static covariation among traits (see also Slatkin 1987).

Evolutionary allometry originates from covariation in the phylogenetic changes of morphometric traits. I do not follow the distinction between data from fossils considered to be a series of ancestors and descendants in a specific evolutionary lineage (evolutionary allometry sensu Gould 1966, 1975) or from contemporaneous species of a clade (Gould's interspecific allometry). The evolutionary processes responsible for morphometric covariation along the branches of a phylogeny are the same, regardless of whether the taxa under study are linked by ancestor–descendant or sister group relationships (the data may have to be adjusted for nonindependence; see, e.g., Felsenstein 1985; Garland et al. 1992). Moreover, I consider evolutionary variation at different taxonomic levels (intraspecific, interspecific) as different expressions of the same phenomenon.

Empirical comparisons have shown that patterns of static and evolutionary allometry can be similar (Gibson et al. 1984; Leamy and Atchley 1984; Klingenberg and Zimmermann 1992). Ontogenetic and evolutionary allometry also can share similar patterns (Wayne 1986a, 1986b; Klingenberg and Zimmermann 1992), which also applies by definition to all cases of ontogenetic scaling (formal comparisons usually are not made, however, because most of these studies deal with a single pair of species). In a comparison of all three levels of allometry in water striders, Klingenberg and Zimmermann (1992) found that patterns of static and ontogenetic allometry were more similar to each other than either was to evolutionary allometry, suggesting differential selection among traits as a possible factor generating the interspecific pattern.

The commonalities between levels of allometry reflect the fact that ontogenetic, static and evolutionary variation are interconnected inextricably because variation in ontogenetic processes supplies static variation upon which natural selection can act. Moreover, the phenotypic consequences of genetic changes by selection or drift depend on the degree to which the developmental processes are canalized (Saunders 1990). Synthesis of current knowledge about development and genetics into heuristic models (e.g., Riska 1986, 1989; Slatkin 1987; Atchley and Hall 1991; Cowley and Atchley 1992) can help un-

derstanding the concerted action of all components of evolutionary processes; moreover, such models can pinpoint critical areas for future research.

Conclusions

The goal of this review is to summarize the concepts of heterochrony and allometry, and to illustrate how they can be applied to investigate evolutionary processes. Throughout, I have emphasized the distinctions among methodological frameworks for both heterochrony and allometry. My intent is not to set up artificial barriers, but to make readers aware of the differences that exist in the literature. These differences, coupled with the use of the same terminology in often incompatible contexts has led to apparent conflicts. These, I think, can be resolved best by examining the logical basis of different research approaches and the meanings of terms in each of them. Recognizing these differences is indispensable for using the complementary strengths of the methods effectively.

The two classical frameworks for analyzing heterochrony differ in the way they compare ontogenies. The clock model of Gould (1977) is a device for displaying entire ontogenies, but the classification of evolutionary changes is based on the differences in size, shape, and age at a standard stage. As this classification is based on entire "constellations" of changes in these three variables, each ancestor–descendant comparison can be unambiguously assigned to one type of heterochronic change, although not all of the possible types are named (Fig. 3-2B). In contrast, the formalism of Alberch et al. (1979) is based on a simplified model that characterizes a developmental process by its onset time, rate, and time of cessation. Each of the terms for heterochronic changes is defined an evolutionary alteration in one of the three model parameters. As these parameters can change simultaneously, combinations of the heterochronic "unit processes" are expected to be the rule rather than the exception.

The framework of Alberch et al. (1979) underwent further change when the terms for heterochronic changes were applied not only to alterations of shape, but also of size variables. Although this practice has extended the usage of the terminology, the modified formalism is consistent with the model of Alberch et al., which was designed explicitly for both size and shape. It reflects a concept of shape as the relative sizes of parts of an organism (which also includes information about their absolute sizes), in contrast to a purely geometric view that focuses exclusively on shape as proportion.

Recognition of the differences helps resolving the controversy about human heterochrony. Most contentious issues can be understood by examining the conceptual basis of different studies. The consensus position emerging from these considerations emphasizes the complex interplay of various processes that is partly independent in different organ systems, rather than the predominance of a single tendency. This consensus originates from a better understanding of the developmental mechanisms involved in evolutionary changes of ontogenies.

Developmental biologists have become increasingly interested in heterochrony as an evolutionary process. Application of the classical frameworks, however, has proven to be difficult because they were designed for the phenomenological description of growth, but do not reflect the interest in control

mechanisms that dominates developmental biology. To incorporate these considerations, and to separate the terms from their historical association with sexual maturation, Raff and Wray (1989) proposed a new set of terms for heterochrony (Table 3-1). Although their classification is based on a model similar to that of Alberch et al (1979), it uses entirely descriptive terms, and it incorporates some assumptions about developmental control via feedback mechanisms.

Time is an essential component of heterochronic analyses. Different measures of extrinsic or intrinsic time will produce different results, and the choice of a metric therefore must be justified for each study. Extrinsic time is used most often, not only because the data are more readily available than for other metrics, but also because extrinsic time provides an unambiguous basis for comparisons and a link to life-history theory, which constitutes the ecological background for the evolution of ontogenies. Several lines of evidence, both theoretical and empirical, demonstrate that size cannot be a substitute for age data; the dimension of time is indispensable for inferring heterochronic processes. Although they cannot be used to identify types of heterochrony, allometric studies in their own right provide important information about the evolution of ontogenies.

Allometry is concerned with the covariation among morphological traits, or with variation in shape associated with differences in size. These two characterizations of allometry reflect differences in concepts of organismal size and shape. Huxley's (1932) approach of allometry, later generalized for multivariate application (Jolicoeur 1963), deals with the relative sizes of parts of organisms, and invokes shape only as an interpretation of the results of the analysis. In contrast, the alternative concept of allometry developed by Gould (1966) and later formalized statistically by Mosimann (1970) specifically refers to shape as a geometric concept. Both approaches use the same basic notion of isometry (although it is formulated in very different ways); isometry therefore provides link between the two schools of thought. Moreover, it is a key concept to relate allometric analyses to the ontogenetic polarity of heterochronic studies. Allometric variation is the product of numerous processes, and studies at different levels of variation can provide insight into the interactions among them.

Allometry and heterochrony are integral parts of the emerging synthesis of evolutionary and developmental biology (e.g., Atkinson 1992; Hall 1992; Raff 1992). A complete understanding of the evolution of ontogenies will require joint studies of variation in morphological traits and timing of developmental events within populations and among related species, combined with information on genetics, developmental mechanisms, life histories, and phylogeny. Although such a research program may seem daunting, the literature reviewed in this article provides examples of empirical studies approaching this goal.

REFERENCES

Alberch P. 1985. Problems with the interpretation of developmental sequences. *Systematic Zoology* 34:46–58.

Alberch P. 1989. The logic of monsters: evidence for internal constraint in development and evolution. In David B, Dommergues J-L, and Laurin B, editors. *Ontogenèse et Évolution. Geobios, mémoire spécial* 12:21–57.

Alberch P. 1991. From genes to phenotype: dynamical systems and evolvability. *Genetica* 84:5–11.

Alberch P, Alberch J. 1981. Heterochronic mechanisms of morphological diversification and evolutionary change in the Neotropical salamander, *Bolitoglossa occidentalis* (Amphibia: Plethodontidae). *Journal of Morphology* 167:249–264.

Alberch P, Gould S J, Oster G F, Wake D B. 1979. Size and shape in ontogeny and phylogeny. *Paleobiology* 5:296–317.

Albrecht G H, Gelvin B R, Hartman S E. 1993. Ratios as a size adjustment in morphometrics. *American Journal of Physical Anthropology* 91:441–468.

Allmon W D. 1994. Patterns and processes of heterochrony in lower Tertiary turritelline gastropods, U.S. Gulf and Atlantic Coastal Plains. *Journal of Paleontology* 68:80–95.

Ambros V. 1988. Genetic basis for heterochronic variation. In McKinney M L, editor. *Heterochrony in Evolution: A Multidisciplinary Approach*. New York: Plenum Press. pp 269–285.

Ambros V. 1989. A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* 57:49–57.

Ambros V, Horvitz H R. 1984. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* 226:409–416.

Ambros V, Moss E G. 1994. Heterochronic genes and the temporal control of *C. elegans* development. *Trends in Genetics* 10:123–127.

Atchley W R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. *American Naturalist* 123:519–540.

Atchley W R. 1987. Developmental quantitative genetics and the evolution of ontogenies. *Evolution* 41:316–330.

Atchley W R, Hall B K. 1991. A model for development and evolution of complex morphological structures. *Biological Reviews* 66:101–157.

Atchley W R, Xu S, Vogl C. 1994. Developmental quantitative genetic models of evolutionary change. *Developmental Genetics* 15:92–103.

Atkinson J W. 1992. Conceptual issues in the reunion of development and evolution. *Synthese* 91:93–110.

Barton A D, Laird A K. 1968. Analysis of allometric and non-allometric differential growth. Growth 33:1–16.

Berge C. 1991. Size- and locomotion-related aspects of hominid and anthropoid pelves: an osteometrical multivariate analysis. *Human Evolution* 6:365–376.

Berge C, Kazmierczak J-B. 1986. Effects of size and locomotor adaptations on the hominid pelvis: evaluation of australopithecine bipedality wit a new multivariate method. *Folia Primatologica* 46:185–204.

Beynon A D, Dean M C. 1988. Distinct dental development patterns in early fossil hominids. *Nature* 335:509–514.

Björklund M. 1993. Phenotypic variation of growth trajectories in finches. *Evolution* 47:1506–1514.

Blackith R E, Davies R G, Moy E A. 1963. A biometric analysis of development in *Dysdercus fasciatus* Sign. (Hemiptera: Pyrrhocoridae). *Growth* 27:317–334.

Blackstone N W. 1986. Relative growth and specific growth rates in crustaceans. *Growth* 50:118–127.

Blackstone N W. 1987a. Allometry and relative growth: pattern and process in evolutionary studies. *Systematic Zoology* 36:76–78.

Blackstone N W. 1987b. Size and Time. *Systematic Zoology* 36:211–215.

Blackstone N W. 1987c. Specific growth rates of parts in a Hermit crab *Pagurus longicarpus*: a reductionist approach to the study of allometry. *Journal of Zoology (London)* 211:531–545.

Blackstone N W, Buss L W. 1992. Treatment with 2,4-dinitrophenol mimics ontogenetic and phylogenetic changes in a hydractiniid hydroid. *Proceedings of the National Academy of Sciences, U.S.A.* 89:4057–4061.

Blackstone N W, Yund P O. 1989. Morphological variation in a colonial marine hydroid: a comparison of size-based and age-based heterochrony. *Paleobiology* 15:1–10.

Blanco M J, Alberch P. 1992. Caenogenesis, developmental variability, and evolution in the carpus and tarsus of the marbled newt *Triturus marmoratus*. *Evolution* 46:677–687.

Boag P T. 1984. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galapagos. *Journal of Zoology, London* 204:413–441.

Bookstein F L. 1989. "Size and shape": a comment on semantics. *Systematic Zoology* 38:173–180.

Bookstein F L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge (UK): Cambridge University Press.

Bookstein F L. 1993. A brief history of the morphometric synthesis. In Marcus L F, Bello E, García-Valdecasas A, editors. *Contributions to Morphometrics*. Madrid: Museo Nacional de Ciencias Naturales. pp 15–40.

Bookstein F L, Chernoff B, Elder R L, Humphries J M, Jr, Smith G R, Strauss R E. 1985. *Morphometrics in Evolutionary Biology: the Geometry of Size and Shape Change, with Examples from Fishes*. Philadelphia: Academy of Natural Sciences of Philadelphia. Special publication 15.

Boughton D A, Collette B B, McCune A R. 1991. Heterochrony in jaw morphology of needlefishes (Teleostei: Belonidae). *Systematic Zoology* 40:329–354.

Brakefield P M, French V. 1995. Eyespot development on butterfly wings: the epidermal response to damage. *Developmental Biology* 168:98–111.

Brown V, Davies R G. 1972. Allometric growth in two species of *Ectobius* (Dictyoptera: Blattidae). *Journal of Zoology, London* 166:97–132.

Bryant H N, Russell A P. 1992. The role of phylogenetic analysis in the inference of unpreserved attributes of extinct taxa. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 337:405–418.

Bryant P J, Simpson P. 1984. Intrinsic and extrinsic control of growth in developing organs. *Quarterly Review of Biology* 59:387–415.

Burnaby T P. 1966. Growth-invariant discriminant functions and generalized distances. *Biometrics* 22:96–110.

Calder, W A, III. 1984. *Size, Function, and Life History*. Cambridge (Mass.): Harvard University Press.

Cane W P. 1993. The ontogeny of postcranial integration in the common tern, *Sterna hirundo*. *Evolution* 47:1138–1151.

Carrier D, Leon L R. 1990. Skeletal growth and function in the California gull (*Larus californicus*). *Journal of Zoology, London* 222:375–389.

Carroll S B. 1994. Developmental regulatory mechanisms in the evolution of insect diversity. *Development* 1994 Supplement:217–223.

Carroll S B, Gates J, Keys D N, Paddock S W, Panganiban G E F, Selegue J E, Williams J A. 1994. Pattern formation and eyespot determination in butterfly wings. *Science* 265:109–114.

Castanet J, Francillon-Vieillot H, Meunier F J, de Ricqlès A. 1993. Bone and individual aging. In Hall B K, editor. *Bone. Vol 7: Bone Growth—B*. Boca Raton: CRC Press. pp 245–283.

Cheverud J M. 1982. Relationships among ontogenetic, static, and evolutionary allometry. *American Journal of Physical Anthropology* 59:139–149.

Cheverud J M, Rutledge J J, Atchley W R. 1983. Quantitative genetics of development: genetic correlations among age-specific trait values and the evolution of ontogeny. *Evolution* 37:895–905.

Coates M I. 1994. The origin of vertebrate limbs. *Development* 1994 Supplement:169–180.

Cock A G. 1966. Genetical aspects of metrical growth and form in animals. *Quarterly Review of Biology* 41:131–190.

Collazo A. 1994. Molecular heterochrony in the pattern of fibronectin expression during gastrulation in amphibians. *Evolution* 48:2037–2045.

Cowley D E, Atchley W R. 1992. Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. *Evolution* 46:495–518.

Creighton G K, Strauss R E. 1986. Comparative patterns of growth and development in cricetine rodents and the evolution of ontogeny. *Evolution* 40:94–106.

Cuzin-Roudy J. 1975. Étude de la variabilité et de l'allométrie de taille chez *Notonecta maculata* Fabricius (insectes, hétéroptères), par les méthodes classiques et par la méthode des composantes principales. *Archives de Zoologie Expérimentale et Générale* 116:173–227.

Cuzin-Roudy J, Laval P. 1975. A canonical discriminant analysis of post-embryonic development in *Notonecta maculata* Frabricius (Insecta: Heteroptera). *Growth* 39:251–280.

Darroch J N, Mosimann J E. 1985. Canonical and principal components of shape. *Biometrika* 72:241–252.

David B. 1989. Jeu en mosaïque des hétérochronies: variation et diversité chez les Pourtalesiidae (échinides abyssaux). In David B, Dommergues J-L, Laurin B, editors. *Ontogenèse et Évolution. Geobios, mémoire spécial* 12:115–131.

David B. 1990. Mosaic pattern of heterochronies: variation and diversity in Poutalesiidae (deep-sea echinoids). *Evolutionary Biology* 24:297–327.

Davies R G, Brown V. 1972. A multivariate analysis of postembryonic growth in two species of *Ectobius* (Dictyoptera: Blattidae). *Journal of Zoology, London* 168:51–79.

Dean M C, Wood B A. 1984. Phylogeny, neoteny and growth of the cranial base in hominoids. *Folia Primatologica* 43:157–180.

de Beer, G. 1958. *Embryos and Ancestors*. Third Edition. Clarendon Press, Oxford.

Dettlaff T A. 1986. The rate of development in poikilothermic animals in astronomical and relative time units. *Journal of Thermal Biology* 11:1–7.

Dettlaff T A, Ignatieva G M, Vassetzky S G. 1987. The problem of time in developmental biology: its study by the use of relative characteristics of development duration. *Soviet Science Review F Physiology and General Biology* 1:1–88.

Dommergues, J-L. 1986. Translation des trajectoires ontogénétiques et résultats hétérochroniques; un modèle. *Comptes Rendus de l'Académie des Sciences de Paris Série II*, 302:1211–1213.

Dommergues J-L, David B, Marchand D. 1986. Les relations ontogenèse-phylogenèse: applications paléontologiques. *Geobios* 19:335–356.

Dommergues J-L, Meister C. 1989. Trajectoires ontogénétiques et hétérochronies complexes chez des ammonites (Harpoceratinae) du jurassique inférieur (Domérien). In David B, Dommergues J-L, and Laurin B, editors. *Ontogenèse et Évolution. Geobios, mémoire spécial* 12:157–166.

Duboule D. 1994. Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development* 1994 Supplement: 135–142.

Elinson R P. 1987. Change in developmental patterns: embryos of amphibians with large eggs. In Raff R A, Raff E C, editors. *Development as an Evolutionary Process*. New York: A. R. Liss. pp 1–21.

Elinson R P. 1989. Egg evolution. In Wake D B, Roth G, editors. *Complex Organismal Functions: Integration and Evolution in Vertebrates*. Chichester: John Wiley, . pp. 251–262.

Emerson S B. 1986. Heterochrony and frogs: the relationship of a life history trait to morphological form. *American Naturalist* 127:167–183.

Emerson S B, Bramble D M. 1993. Scaling, allometry, and skull design. In Hanken J, Hall B K editors. *The Skull*. Volume. 3. Chicago: University of Chicago Press. pp 384–421.

Emlen D J. 1994. Environmental control of horn length dimorphism in the beetle *Onthophagus acuminatus* (Coleoptera: Scarabaeidae). *Proceedings of the Royal Society of London B Biological Sciences* 256:131–136.

Falconer D S. 1989. *Introduction to Quantitative Genetics*. Third Edition. Essex (UK): Longman.

Felsenstein J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Flessa K W, Bray R G. 1977. On the measurement of size-independent morphological variability: an example using successive populations of a Devonian spiriferid brachiopod. *Paleobiology* 3:350–359.

Freeman G, Lundelius J W. 1992. Evolutionary implications of the mode of D quadrant specification in coelomates with spiral cleavage. *Journal of Evolutionary Biology* 5:205–247.

French V, Brakefield P M. 1992. The development of eyespot patterns on butterfly wings: morphogen sources or sinks? *Development* 116:103–109.

French V, Brakefield P M. 1995. Eyespot development on butterfly wings: the focal signal. *Developmental Biology* 168:112–123.

Gallardo R, Dominguez E, Muñoz J M. 1993. The heterochronic origin of the cleistogamous flower in *Astragalus cymbicarpos* (Fabaceae). *American Journal of Botany* 80:814–823.

Garcia-Bellido A, de Celis J F. 1992. Developmental genetics of the venation pattern of *Drosophila*. *Annual Review of Genetics* 26:277–304.

Garland T, Jr, Harvey P H, Ives A R. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Systematic Biology* 41:18–32.

German R Z, Meyers L L. 1989a. The role of time and size in ontogenetic allometry: I. Review. *Growth Development and Aging* 53:101–106.

German R Z, Meyers L L. 1989b. The role of time and size in ontogenetic allometry: II. An empirical study of human growth. *Growth Development and Aging* 53:107–115.

Gibson A R, Baker A J, Moeed A. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): an application of the jackknife to principal component analysis. *Systematic Zoology* 33:408–421.

Gilbert S F, Opitz J M, Raff R A. 1996. Resynthesizing evolutionary and developmental biology. *Developmental Biology* 173:357–372.

Godfrey L R, Sutherland M R. 1995a. Flawed inference: why size-based tests of heterochronic processes do not work. *Journal of Theoretical Biology* 172:43–61.

Godfrey L R, Sutherland M R. 1995b. What's growth got to do with it? Process and product in the evolution of ontogeny. *Journal of Human Evolution* 29:405–431.

Godfrey L R, Sutherland M R. 1996. Paradox of peramorphic paedomorphosis: heterochrony and human evolution. *American Journal of Physical Anthropology* 99:17–42.

Gould S J. 1966. Allometry and size in ontogeny and phylogeny. *Biological Reviews* 41:587–640.

Gould S J. 1975. Allometry in primates, with emphasis on scaling and the evolution of the brain. *Contributions to Primatology* 5:244–292.

Gould S J. 1977. *Ontogeny and Phylogeny*. Cambridge, Mass.: Harvard University Press.

Gould S J. 1982. Change in developmental timing as a mechanism of macroevolution. In Bonner J T, editor. *Evolution and Development*. Berlin: Springer-Verlag.

Gould S J. 1992. Heterochrony. In Keller E F, Lloyd E A, editors. *Keywords in Evolutionary Biology*. Cambridge, Mass.: Harvard University Press.

Guerrant E O, Jr. 1988. Heterochrony in plants: the intersection of evolutionary ecology and ontogeny. In McKinney M L, editor. *Heterochrony in Evolution: a Multidisciplinary Approach*. New York: Plenum Press. pp 111–133.

Guihard-Costa A-M. 1991. Vitesse de croissance foetale: une nouvelle approche méthodologique. *Comptes Rendus de l'Acadademie des Sciences de Paris Série III Sciences de la Vie* 312:403–405.

Hall B K. 1984. Developmental processes underlying heterochrony as an evolutionary mechanism. *Canadian Journal of Zoology* 62:1–7.

Hall B K. 1990. Heterochrony in vertebrate development. *Seminars in Developmental Biology* 1:237–243.

Hall B K. 1992. *Evolutionary Developmental Biology*. Chapman and Hall, London.

Hall B K, Miyake T. 1995. How do embryos measure time? In McNamara K J, editor. Evolutionary Change and Heterochrony. Chichester: John Wiley & Sons. pp 3–20.

Hanken J. 1993. Model systems versus outgroups: alternative approaches to the study of head development and evolution. *American Zoologist* 33:448–456.

Healy M J R, Tanner J M. 1981. Size and shape in relation to growth and form. *Symposia of the Zoological Society of London* 46:19–35.

Henry J J, Raff R A. 1990. Evolutionary change in the process of dorsoventral axis determination in the direct developing sea urchin, *Heliocidaris erythrogramma*. *Developmental Biology* 141:55–69.

Herring S W. 1993. Epigenetic and functional influences on skull growth. In Hanken J, Hall B K, editors. *The Skull. Volume 1: Development*. Chicago: University of Chicago Press. pp 153–206.

Hinchliffe R J. 1994. Evolutionary developmental biology of the tetrapod limb. *Development* 1994 Supplement:163–168.

Huxley J S. 1924. Constant differential growth-ratios and their significance. *Nature* 114:895–896.

Huxley J S. 1932. *Problems of Relative Growth*. London: Methuen. Reprinted, 1993, Baltimore: Johns Hopkins University Press.

Ishikawa A, Namikawa T. 1987. Postnatal growth and development in laboratory strains of large and small musk shrews (*Suncus murinus*). *Journal of Mammalogy* 68:766–774.

Jolicoeur P. 1963. The multivariate generalization of the allometry equation. *Biometrics* 19:497–499.

Jolicoeur P, Mosimann J E. 1960. Size and shape variation in the painted turtle: a principal component analysis. *Growth* 24:339–354.

Jones C S. 1992. Comparative ontogeny of a wild cucurbit and its derived cultivar. *Evolution* 46:1827–1847.

Jones C S. 1993. Heterochrony and heteroblastic leaf development in two subspecies of *Cucurbita argyrosperma* (Cucurbitaceae). *American Journal of Botany* 80:778–795.

Jones D S. 1988. Sclerochronology and the size versus age problem. In McKinney M L, editor. *Heterochrony in Evolution: a Multidisciplinary Approach*. New York: Plenum Press. pp 93–108.

Jungers W L, Falsetti A B, Wall C E. 1995. Shape, relative size, and size-adjustments in morphometrics. *Yearbook of Physical Anthropology* 38:137–161.

Kampny C M, Dickinson T A, Dengler N G. 1993. Quantitative comparison of floral development in *Veronica chamaedrys* and *Veronicastrum virginicum* (Scrophulariaceae). *American Journal of Botany* 80:449–460.

Kampny C M, Dickinson T A, Dengler N G. 1994. Quantitative floral development in *Pseudolysimachion* (Scrophulariaceae): intraspecific variation and comparison with *Veronica* and *Veronicastrum*. *American Journal of Botany* 81:1343–1353.

Katz M J. 1980. Allometry formula: a cellular model. *Growth* 44:89–96.

Kazmierczak J B. 1985. Analyse logarithmique: deux exemples d'application. *Revue de Statistique Appliquée* 33:13–24.

Keene H J. 1982. The morphogenetic triangle: a new conceptual tool for application to problems in dental morphogenesis. *American Journal of Physical Anthropology* 59:281–287.

Keene H J. 1991. On heterochrony in heterodonty: A review of some problems of tooth morphogenesis and evolution. *Yearbook of Physical Anthropology* 34:251–282.

Kirkpatrick M, Lofsvold D. 1989. The evolution of growth trajectories and other complex quantitative characters. *Genome* 31:778–783.

Kirkpatrick M, Lofsvold D. 1992. Measuring selection and constraint in the evolution of growth. *Evolution* 46:954–971.

Klingenberg C P. 1996a. Multivariate allometry. In Marcus L F, Corti M, Loy A, Naylor G, Slice, D E, editors. *Advances in Morphometrics*. New York: Plenum Press. pp 23–49.

Klingenberg C P. 1996b. Individual variation of ontogenies: A longitudinal study of growth and timing. *Submitted to Evolution*.

Klingenberg C P, Ekau W. 1996. A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). *Biological Journal of the Linnean Society*. In press.

Klingenberg C P, Neuenschwander B E, Flury B D. 1996. Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components. *Systematic Biology*. In press.

Klingenberg C P, Spence J R. 1993. Heterochrony and allometry: lessons from the water strider genus *Limnoporus*. *Evolution* 47:1834–1853.

Klingenberg C P, Spence J R. 1996. On the role of body size for life history evolution. *Ecological Entomology*. In press.

Klingenberg C P, Zimmermann M. 1992. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of water striders. *American Naturalist* 140:601–620.

Laird A K. 1965. Dynamics of relative growth. *Growth* 29:249–263.

Laird A K, Barton A D, Tyler S A. 1968. Growth and time: an interpretation of allometry. Growth 32:347–354.

Lande R. 1985. Genetic and evolutionary aspects of allometry. In Jungers W L, editor. *Size and Scaling in Primate Biology*. New York: Plenum Press. pp 21–32.

Leamy L, Atchley W. 1984. Static and evolutionary allometry of osteometric traits in selected lines of rats. *Evolution* 38:47–54.

Leamy L, Bradley D. 1982. Static and growth allometry of morphometric traits in randombred house mice. *Evolution* 36:1200–1212.

Leamy L, Cheverud J M. 1984. Quantitative genetics and the evolution of ontogeny. II. Genetic and environmental correlations among age-specific characters in randombred house mice. *Growth* 48:339–353.

Leclerc R F, Regier J C. 1990. Heterochrony in insect development and evolution. *Seminars in Developmental Biology* 1:271–279.

Leigh S R, Shea B T. 1996. Ontogeny of body size variation in African apes. *American Journal of Physical Anthropology* 99:43–65.

Lele S, Richtsmeier J T. 1991. Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology* 86:415–427.

Lessa E P, Patton J L. 1989. Structural constraints, recurrent shapes, and allometry in pocket gophers (genus *Thomomys*). *Biological Journal of the Linnean Society* 36:349–363.

Liem K F. 1991. Toward a new morphology: pluralism in research and education. *American Zoologist* 31:759–767.

Livezey B C. 1992. Flightlessness in the Galápagos cormorant (*Compsohalieus* [*Nannopterum*] *harrisi*): heterochrony, giantism and specialization. *Zoological Journal of the Linnean Society* 105:155–224.

Livezey B C. 1995. Heterochrony and the evolution of avian flightlessness. In McNamara K J, editor. *Evolutionary Change and Heterochrony*. Chichester: John Wiley & Sons. pp 169–193.

Lord E M, Hill J P. 1987. Evidence for heterochrony in the evolution of plant form. In Raff R A, Raff E C, editors. *Development as an Evolutionary Process*. New York: Alan R. Liss, Inc. pp 47–70.

Lynch M. 1988. Path analysis of ontogenetic data. In Ebenman B, Persson L, editors. *Size-Structured Populations*. Berlin: Springer-Verlag. pp 29–46.

Maddison W P, Maddison D R. 1992. *Mac Clade: Analysis of Phylogeny and Character Evolution*. Sunderland: Sinauer Associates.

Marshall W A, Tanner J M. 1986. Puberty. In Falkner F, Tanner J M, editors. *Human Growth: A Comprehensive Treatise*. Vol. 2. Second edition. New York: Plenum Press. pp 171–209.

Matsuda R. 1961. Studies of relative growth in Gerridae (Hemiptera-Heteroptera), I–III. *Annals of the Entomological Society of America* 54:578–598.

Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell, J, Goodwin B, Lande R, Raup D, Wolpert L. 1985. Developmental constraints and evolution. *Quarterly Review of Biology* 60:265–287.

McKinney M L. 1984. Allometry and heterochrony in an Eocene echinoid lineage: morphological change as a by-product of size selection. *Paleobiology* 10:407–419.

McKinney M L. 1986. Ecological causation of heterochrony: a test and implications for evolutionary theory. *Paleobiology* 12:282–289.

McKinney M L. 1988. Classifying heterochrony: allometry, size, and time. In McKinney M L editor. *Heterochrony in Evolution: A Multidisciplinary Approach*. New York: Plenum Press. pp 17–34.

McKinney M L, Gittleman J L. 1995. Ontogeny and phylogeny: tinkering with covariation in life history, morphology, and behaviour. In McNamara K J, editor. *Evolutionary Change and Heterochrony*. Chichester: John Wiley & Sons. pp 21–47.

McKinney M L, McNamara K J. 1991. *Heterochrony: the Evolution of Ontogeny*. New York: Plenum Press.

McLellan T. 1993. The roles of heterochrony and heteroblasty in the diversification of leaf shapes in *Begonia dregei* (Begoniaceae). *American Journal of Botany* 80:796–804.

McNamara K J. 1986. A guide to the nomenclature of heterochrony. *Journal of Paleontology* 60:4–13.

McNamara K J. 1988. The abundance of heterochrony in the fossil record. In McKinney M L, editor. *Heterochrony in Evolution: A Multidisciplinary Approach*. New York: Plenum Press. pp 287–325.

Merimee T J, Zapf J, Hewlett B, Cavalli-Sforza L L. 1987. Insulin-life growth factors in pygmies: the role of puberty in determining final stature. *New England Journal of Medicine* 316:906–911.

Montagu A. 1989. *Growing young*. 2nd edition. Granby, Mass.: Bergin & Garvey.

Morgan B A, Tabin C. 1994. Hox genes and growth: early and late roles in limb bud morphogenesis. *Development* 1994 Supplement:181–186.

Mosimann J E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. *Journal of the American Statistical Association* 65:930–945.

Mosimann J E, James F C. 1979. New statistical methods for allometry with application to Florida red-winged blackbirds. *Evolution* 33:444–459.

Nijhout H F. 1980. Pattern formation on lepidopteran wings: determination of an eyespot. *Developmental Biology* 80:267–274.

Nijhout H F. 1985. Cautery-induced colour patterns in *Precis coenia* (Lepidoptera: Nymphalidae). *Journal of Embryology and Experimental Morphology* 86:191–203.

Nijhout H F. 1990a. A comprehensive model for colour pattern formation in butterflies. *Proceedings of the Royal Society of London B Biological Sciences* 239:81–113.

Nijhout H F. 1990b. Metaphors and the role of genes in development. *BioEssays* 12:441–446.

Nijhout H F. 1991. *The Development and Evolution of Butterfly Wing Patterns*. Washington: Smithsonian Institution Press.

Nijhout H F. 1994a. *Insect Hormones*. Princeton: Princeton University Press.

Nijhout H F. 1994b. Symmetry systems and compartments in Lepidopteran wings: the evolution of a patterning mechanism. *Development* 1994 Supplement:225–233.

Novák, V J A. 1966. *Insect Hormones*. London: Methuen.

Okazaki K, Dan K. 1954. The metamorphosis of partial larvae of *Peronella japonica* Mortensen, a sand dollar. *Biological Bulletin* 106:83–99.

Oster G F, Shubin N, Murray J D, Alberch P. 1988. Evolution and morphogenetic rules: The shape of the vertebrate limb in ontogeny and phylogeny. *Evolution* 45:862–884.

Parks A L, Parr B A, Chin J-E, Leaf D S, Raff R A. 1988. Molecular analysis of heterochronic changes in the evolution of direct developing sea urchins. *Journal of Evolutionary Biology* 1:27–44.

Powell J R, DeSalle R. 1995. *Drosophila* molecular phylogenies and their uses. *Evolutionary Biology* 28:87–138.

Pruess K P. 1983. Day-degree methods for pest management. *Environmental Entomology* 12:613–619.

Raff R A. 1987. Constraint, flexibility, and phylogenetic history in the evolution of direct development in sea urchins. *Developmental Biology* 119:6–19.

Raff R A. 1992. Evolution of developmental decisions and morphogenesis: the view from two camps. *Development* 1992 Supplement:15–22.

Raff R A, Wray G A. 1989. Heterochrony: developmental mechanisms and evolutionary results. *Journal of Evolutionary Biology* 2:409–434.

Ravosa M J, Meyers D M, Glander K E. 1993. Relative growth of the limbs and trunk in sifakas: heterochronic, ecological, and functional considerations. *American Journal of Physical Anthropology* 92:499–520.

Reeve E C R, Huxley J S. 1945. Some problems in the study of allometric growth. In Le Gros Clark W E, Medawar P, editors. *Essays on Growth and Form Presented to D'Arcy Wentworth Thompson*. Oxford: Clarendon Press. pp 121–156.

Regier J C, Vlahos N S. 1988. Heterochrony and the introduction of novel modes of morphogenesis during the evolution of moth choriogenesis. *Journal of Molecular Evolution* 28:19–31.

Reilly S M. 1994. The ecological morphology of metamorphosis: heterochrony and the evolution of feeding mechanisms in salamanders. In Wainwright P C, Reilly S M, editors. *Ecological Morphology: Integrative Organismal Biology*. Chicago: University of Chicago Press. pp 319–338.

Reilly S M, Wiley E O, Meinhardt D J. 1996. An integrative approach to heterochrony: the distinction between interspecific and intraspecific phenomena. *Biological Journal of the Linnean Society*. In press.

Reiss J O. 1989. The meaning of developmental time: a metric for comparative embryology. *American Naturalist* 134:170–189.

Reiss M J. 1989. *The Allometry of Growth and Reproduction*. Cambridge (UK): Cambridge University Press.

Richardson M K. 1995. Heterochrony and the phylotypic period. *Developmental Biology* 172:412–421.

Richtsmeier J T, Corner B D, Grausz H M, Cheverud J M, Danahey S E. 1993. The role of postnatal growth pattern in the production of facial morphology. *Systematic Biology* 42:307–330.

Richtsmeier J T, Lele S. 1993. A coordinate-free approach to the analysis of growth patterns: Models and theoretical considerations. *Biological Reviews* 68:381–411.

Riska B. 1986. Some models for development, growth, and morphometric correlation. *Evolution* 40: 1303–1311.

Riska B. 1989. Composite traits, selection response, and evolution. *Evolution* 43:1172–1191.

Riska B, Atchley W R, Rutledge J J. 1984. A genetic analysis of targeted growth in mice. *Genetics* 107:79–101.

Ritterbusch A. 1990. The measure of biological age in plant modular systems. *Acta Biotheoretica* 38:113–124.

Rohlf F J. 1990. Morphometrics. *Annual Review of Ecology and Systematics* 21:299–316.

Rohlf F J, Bookstein F L. 1990. *Proceedings of the Michigan Morphometrics Workshop*. Ann Arbor: University of Michigan Museum of Zoology. Special publication no. 2.

Rohlf F J, Bookstein F L. 1987. A comment on shearing as a method for "size correction." *Systematic Zoology* 36:356–367.

Rohlf F J, Slice D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* 39:40–59.

Ross C, Henneberg M. 1995. Basicranial flexion, relative brain size, and facial kyphosis in *Homo sapiens* and some fossil hominids. *American Journal of Physical Anthropology* 98:575–593.

Saunders P T. 1990. The epigenetic landscape and evolution. *Biological Journal of the Linnean Society* 39:125–134.

Shea B T. 1983a. Allometry and heterochrony in the African apes. *American Journal of Physical Anthropology* 62:275–289.

Shea B T. 1983b. Paedomorphosis and neoteny in the pygmy chimpanzee. *Science* 222:521-522.

Shea B T. 1985. Bivariate and multivariate growth allometry: statistical and biological considerations. *Journal of Zoology (London)* 206:367–390.

Shea B T. 1988. Heterochrony in primates. In McKinney M L, editor. *Heterochrony in Evolution*. New York: Plenum Press. pp 237–266.

Shea B T. 1989. Heterochrony in human evolution: the case for neoteny reconsidered. *Yearbook of Physical Anthropology* 32:69–101.

Shea B T. 1992a. Developmental perspective on size change and allometry in evolution. *Evolutionary Anthropology* 1:125–134.

Shea B T. 1992b. Ontogenetic scaling of skeletal proportions in the talapoin monkey. *Journal of Human Evolution* 23:283–307.

Shea B T, Hammer R E, Brinster R L, Ravosa M R. 1990. Relative growth of the skull and postcranium in giant transgenic mice. *Genetical Research (Cambridge)* 56:21–34

Shubin N. 1995. The evolution of paired fins and the origin of tetrapod limbs: phylogenetic and transformational approaches. *Evolutionary Biology* 28:39–86.

Shubin N H, Alberch P. 1986. A morphogenetic approach to the origin and basic organization of the tetrapod limb. *Evolutionary Biology* 20:319–387.

Shubin N, Wake D B, Crawford A J. 1995. Morphological variation in the limbs of *Taricha granulosa* (Caudata: Salamandridae): evolutionary and phylogenetic implications. *Evolution* 49:874–884.

Sinervo B. 1993. The effect of offspring size on physiology and life history. *BioScience* 43:210–218.

Sinervo B, Doyle R W. 1990. Life-history analysis in "physiological" compared with "sidereal" time: an example with an amphipod (*Gammarus lawrencianus*) in a varying environment. *Marine Biology* 107:129–139.

Sinervo B, McEdward L R. 1988. Developmental consequences of an evolutionary change in egg size: An experimental test. *Evolution* 42:885–899.

Slatkin M. 1987. Quantitative genetics of heterochrony. *Evolution* 41:799–811.

Smith B H, Tompkins R L. 1995. Toward a life history of the Hominidae. *Annual Review of Anthropology* 24:257–279.

Smith M M. 1995. Heterochrony in the evolution of enamel in vertebrates. In McNamara K J, editor. *Evolutionary Change and Heterochrony*. Chichester: John Wiley & Sons. pp 125–150.

Somers K M. 1989. Allometry, isometry and shape in principal components analysis. *Systematic Zoology* 38:169–173.

Sprent P. 1972. The mathematics of size and shape. *Biometrics* 28:23–37.

Starck J M. 1993. Evolution of avian ontogenies. *Current Ornithology* 10:275–366.

Strathmann R R. 1978. The evolution and loss of feeding larval stages of marine invertebrates. *Evolution* 32:894–906.

Strathmann R R. 1985. Feeding and nonfeeding larval development and life-history evolution in marine invertebrates. *Annual Review of Ecology and Systematics* 16:339–361.

Strathmann R R, Fenaux L, Strathmann M F. 1992. Heterochronic developmental plasticity in larval sea urchins and its implications for evolution of nonfeeding larvae. *Evolution* 46:972–986.

Strauss R E. 1987. On allometry and relative growth in evolutionary studies. *Systematic Zoology* 36:72–75.

Strauss R E. 1990. Heterochronic variation in the developmental timing of cranial ossification in poecilid fishes (Cyprinodontiformes). *Evolution* 44:1558–1567.

Strauss R E, Fuiman L A. 1985. Quantitative comparisons of body form and allometry in larval and adult Pacific sculpins (Teleostei: Cottidae). *Canadian Journal of Zoology* 63:1582–1589.

Sturtevant M A, Bier E. 1995. Analysis of the genetic hierarchy guiding wing vein development in *Drosophila*. *Development* 121:785–801.

Taylor F. 1981. Ecology and evolution of physiological time in insects. *American Naturalist* 117:1–23.

Vrba E S. 1994. An hypothesis of heterochrony in response to climatic cooling and its relevance to early hominid evolution. In Ciochon R, Corruccini R, editors. *Integrative Paths to the Past: Paleanthropological Advances in Honour of F. Clark Howell*. Englewood Cliffs: Prentice-Hall. pp 345–376.

Vrba E S, Vaisnys J R, Gatesy J E, DeSalle R, Wei K-Y. 1994. Analysis of paedomorphosis using allometric characters: the example of Reduncini antelopes (Bovidae, Mammalia). *Systematic Biology* 43:92–116.

Wake D B. 1991. Homoplasy: the result of natural selection, or evidence of design limitations? *American Naturalist* 138:543–567.

Wake D B, Larson A. 1987. Multidimensional analysis of an evolving lineage. *Science* 238:42–48.

Walker J A. 1993. Ontogenetic allometry of threespine stickleback body form using landmark-based morphometrics. In Marcus L F, Bello E, García-Valdecasas A, editors. *Contributions to Morphometrics*. Madrid: Museo Nacional de Ciencias Naturales. pp 193–214.

Wayne R K. 1986a. Cranial morphology of domestic and wild canids: the influence of development on morphological change. *Evolution* 40:243–261.

Wayne R K. 1986b. Limb morphology of domestic and wild canids: the influence of development on morphologic change. *Journal of Morphology* 187:301–319.

Wei K–Y. 1994. Allometric heterochrony in the Pliocene–Pleistocene planktic foraminiferal clade *Globoconella*. *Paleobiology* 20:66–84.

Wheeler D E. 1990. The developmental basis of worker polymorphism in fire ants. *Journal of Insect Physiology* 36:315–322.

Wheeler D E. 1991. The developmental basis of worker caste polymorphism in ants. *American Naturalist* 138:1218–1238.

Whiteman H H. 1994. Evolution of facultative paedomorphosis in salamanders. *Quarterly Review of Biology* 69:205–221.

Wilson S R, Loesch D Z. 1989. Principal component analysis of shape variables in adult individuals. *Annals of Human Biology* 16:361–368.

Wiltshire R J E, Murfet I C, Reid J B. 1994. The genetic control of heterochrony: evidence from developmental mutants of *Pisum sativum* L. *Journal of Evolutionary Biology* 7:447–465.

Winterbottom R. 1990. The *Trimmatom nanus* species complex (Actinopterygii, Gobiidae): phylogeny and progenetic heterochrony. *Systematic Zoology* 39:253–265.

Wood B. 1996. Hominid palaeobiology: have studies of comparative development come of age? *American Journal of Physical Anthropology* 99:9–15.

Wray G A. 1992 The evolution of larval morphology during the post-Paleozoic radiation of echinoids. *Paleobiology* 18:258–287.

Wray G A. 1995. Causes and consequences of heterochrony in early echinoderm development. In McNamara K J, editor. *Evolutionary Change and Heterochrony*. Chichester: John Wiley & Sons. pp 197–223.

Wray G A, Bely A E. 1994. The evolution of echinoderm development is driven by several distinct factors. *Development* 1994 Supplement:97–106.

Wray, G A, McClay D R. 1989. Molecular heterochronies and heterotopies in early echinoid development. *Evolution* 43:803–813.

Wray G A, Raff R A. 1990. Pattern and process heterochronies in the early development of sea urchins. *Seminars in Developmental Biology* 1:245–251.

Wray G A, Raff R A. 1991a. The evolution of developmental strategy in marine invertebrates. *Trends in Ecology and Evolution* 6:45-50.

Wray G A, Raff R A. 1991b. Rapid evolution of gastrulation mechanisms in a sea urchin with lecithotrophic larvae. *Evolution* 45:1741–1750.

Yamashita K, Tanaka S, Iwasawa H. 1991. Is artificial induction of neoteny in hynobiid salamanders possible? Effects of low temperature on the appearance and dynamics of gonadotropin-producing cells in *Hynobius retardatus* and *Hynobius nigrescens*. *Science Reports of Niigata University, Series D (Biology)* 28:1–10.

Yoccoz N G. 1993. Morphométrie et analyses multidimensionelles: une revue des méthodes séparant taille et forme. In Lebreton J-D, Asselain B, editors. *Biométrie et Environnement*. Paris: Masson.

Zelditch M L. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. *Evolution* 42:28–41.

Zelditch M L, Bookstein F L, Lundrigan B L. 1992. Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer*. *Evolution* 46:1164–1180.

Zelditch M L, Bookstein F L, Lundrigan B L. 1993. The ontogenetic complexity of developmental constraints. *Journal of Evolutionary Biology* 6:621–641.

Zelditch M L, Carmichael A C. 1989. Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon fulviventer*. *Evolution* 43:814–824.

Zelditch M L, Fink W L. 1995. Allometry and developmental integration of body growth in a piranha, *Pygocentrus nattereri* (Teleostei: Ostariophysi). *Journal of Morphology* 223:341–355.

4. Heterochrony and Allometry in the Water Strider Genus *Limnoporus*·*

INTRODUCTION

Developmental processes are the proximate origin of all variation in structural characters. Their integration into evolutionary theory under the heading of heterochrony has contributed significantly to our understanding of morphological evolution. The study of heterochrony, now defined as evolutionary change in rates and timing of developmental processes, has a long history of debate and confusion. Gould (1977) reviewed this debate and presented his "clock model" for heterochronic changes as a new conceptual scheme for comparisons of ancestral and descendant ontogenies with respect to size, shape, and age. A slightly different formalism for analyzing and classifying heterochronic phenomena was proposed by Alberch et al. (1979), based on the assumption that a developmental process can be characterized by three parameters: the onset time, the time of completion, and the rate of the process. Recent work has stressed that phylogenetic information is essential to determine the direction of heterochronic changes (Fink, 1982, 1988). The general approach advocated by Alberch et al. (1979), coupled with the use of phylogenetic information, has been adopted in most recent works on heterochrony (McNamara, 1986; McKinney, 1988; Raff and Wray, 1989; McKinney and McNamara, 1991).

Despite the general acceptance of the theoretical framework, its implementations differ widely, and this has led to new confusion in terminology and underlying concepts, and to contradictory interpretations of the same evolutionary events (e.g., the controversy about the relative role of neoteny or hypermorphosis in human evolution [McKinney and McNamara, 1991]). Partly, these problems originate from differences between the conceptual frameworks used by Gould (1977) and Alberch et al. (1979). Other difficulties arise when investigators fail to realize that different organs may follow different heterochronic trends within a given evolutionary lineage, and that several heterochronic changes may affect each trait simultaneously. A clear conceptual separation of patterns and processes involved in heterochronic phenomena can help resolve these difficulties.

A problem encountered by many empirical studies of heterochrony is the absence of age data. In many such cases, measures of size have been substituted for age (e.g., Alberch and Alberch, 1981; McNamara, 1988; Winterbottom, 1990; Boughton et al., 1991), leading to the additional concept of "size-based heterochrony" or "allometric heterochrony" for comparisons of ancestral and descendant allometries (McKinney, 1986, 1988; McKinney and McNamara, 1991). In this paper, we examine the conceptual relationship between heterochrony and allometry, and illustrate the resulting views with a case study of the water strider genus *Limnoporus*.

---

* A version of this chapter has been published. Klingenberg and Spence 1993. Evolution. 47: 1834–1853.

HETEROCHRONY AND ALLOMETRY

Alberch et al. (1979) proposed a framework for analyzing heterochrony by comparing the ontogenies of ancestral and descendant species. The formalism is based on a phenomenological model of a developmental process, which is characterized by three parameters: growth rate, time of onset of growth, and time of its termination. Depending on the parameters that are increased or decreased by evolutionary change, adults of a descendant species may resemble juvenile forms of its ancestor (paedomorphosis) or, conversely, the descendant's development may go beyond the ancestral adult condition (peramorphosis; Fig. 3-3). Alberch et al. (1979 p. 304) treated growth rates of size and shape separately. They classified increases and decreases of the growth rate for shape as acceleration and neoteny, respectively, and increases and decreases of the growth rate for size as proportioned giantism and proportioned dwarfism, respectively. More recent work (McKinney, 1988; McKinney and McNamara, 1991) mostly abandoned this separation, applying the formalism (Fig. 3-3) to any measure of size or shape, or to measurements of a single organ. This practice is consistent with the model for developmental processes, which Alberch et al. (1979 p. 301) explicitly proposed for either size or shape. Raff and Wray (1989) proposed a similar formalism, but used different terms.

From this process-based view of heterochrony, it is clear that the different kinds of heterochronic changes are not mutually exclusive (except those changing the same parameter in opposite directions). The formalism of Alberch et al. (1979) thus cannot be a rigid classification system. Therefore, in a particular case, the question is not whether there is, e.g., either neoteny or hypermorphosis, but what the relative importance of these processes is for the observed evolutionary change. This approach is especially suitable for lineages where several heterochronic processes may act simultaneously (Dommergues et al., 1986) or sequentially ("sequential heterochrony"), or where different organs may display different heterochronic trends ("dissociated heterochrony") rather than one common heterochronic process affecting all aspects of form ("global heterochrony"; McKinney and McNamara, 1991).

Unlike heterochrony, allometry does not refer explicitly to a measure of age, but deals only with the space spanned by the morphological characters. Allometry is concerned with the associations among a number of morphometric traits, or between a trait and a measure of overall size (e.g., Cock, 1966; Gould, 1966; Shea, 1985; Klingenberg and Zimmermann, 1992b). In ontogenetic allometry, which characterizes the trait covariation in samples of organisms that vary in age, there is an implicit relationship to age, as size usually increases with age. This relationship, however, is highly non-linear in many cases. Patterns of ontogenetic allometry reflect the relative growth of the traits, and therefore they may be altered by heterochronic changes that modify growth dynamics.

McKinney (1986) proposed to extend the terminology of heterochronic changes to allometry, replacing age by a measure of size as a reference dimension (see also McKinney, 1988; Lessa and Patton, 1989; McKinney and McNamara, 1991). In a bivariate allometric plot of a trait against size, an extension of the ancestral allometric trajectory to larger sizes in a descendant species is called "allometric hypermorphosis," whereas termination of growth at smaller sizes is "allometric progenesis". Increase or decrease in slope is "allometric acceleration" or "allometric neoteny", respectively, and a larger or smaller *y*-inter-

cept is termed "allometric pre-displacement" or "allometric post-displacement", respectively.

To examine the relationship between age-based heterochrony and allometry, we examine how heterochronic changes in a trait, size, or both affect bivariate allometric plots (Fig. 3-6). For simplicity, we assume that log-transformed measurements of both the trait of interest and overall size follow Figure 3-3 in terms of the ancestral growth dynamics and heterochronic changes. More realistic assumptions would produce more complex allometric plots, but would not change our main conclusions.

In the resulting allometric plots (Fig. 3-6), we find that the underlying heterochronic processes correspond to McKinney's patterns of allometric heterochrony only if certain conditions are met, which differ between the various kinds of heterochrony. For neoteny and acceleration, the expected patterns only result if the trait alone is affected, and for progenesis and hypermorphosis if both the trait and size are affected by the heterochronic change. Pre-displacement and post-displacement generate biphasic allometries, with only one of the two measurements growing during the first phase of the descendant ontogeny. The patterns are consistent with McKinney's scheme of allometric heterochrony only if this phase is disregarded, e.g., by assuming it is outside the time interval for which data are available (e.g., during embryonic development), and if pre- or post-displacement affects only the trait but not size. Progenesis and hypermorphosis for either the trait or size alone generate biphasic allometries with a change in slope late in descendant ontogeny, when data are likely to be available. Moreover, none of the bivariate allometric patterns of McKinney's scheme is unique to a particular heterochronic process, and all changes affecting both the trait and size simultaneously (global heterochrony) are classified as "allometric progenesis" or "allometric hypermorphosis". McKinney (1988; see also McKinney and McNamara, 1991) warned that allometric heterochrony assumes that the growth dynamics of overall size are the same in ancestor and descendant (i.e., that the heterochronic change affects the trait alone). As Figure 3-6 shows, however, this is only true for neoteny and acceleration. As there is no consistent set of conditions that produces the expected allometric patterns for all types of heterochronic changes, we reject the concept of allometric heterochrony. Changes in allometric patterns are the result of changes in ontogeny by heterochronic processes, but the latter cannot be inferred simply from knowledge of the former (see also Blackstone and Yund, 1989). Allometric terminology should reflect this conceptual distinction between patterns and processes, and not use terms associated with heterochrony.

Changes in allometric patterns, both bivariate and multivariate, can be described as changes in the direction of ontogenetic trajectories, lateral transpositions of entire trajectories, and shifts in the positions of particular life-history stages along trajectories. Directions of allometric growth trajectories reflect the relative magnitudes of specific growth rates of the traits studied (Shea, 1985; Blackstone, 1987a). Changes in allometric slopes cannot automatically be attributed to neoteny or acceleration, however, because temporal displacement of growth curves may have the same effect, at least for some growth functions such as the Gompertz curve (Laird et al., 1968). Longitudinal shifts of developmental stages along an ancestral allometry can be produced by any global heterochrony, whereas lateral transpositions of trajectories (differences in allometric

intercepts) may result from any heterochronic process acting at younger ages than those of the organisms included in the study.

In a multivariate context, allometry can help to identify patterns of covariation among traits, and to find alterations of ontogenetic trajectories caused by heterochronic processes. In this paper, we integrate allometry and heterochrony in a case study of the water strider genus *Limnoporus*. We use multivariate techniques to characterize allometric trajectories, and to define a measure of overall size for the analysis of heterochrony. In addition to the morphometric data, we base our interpretations of heterochrony on age data and a reconstructed phylogeny of the genus.

Materials and Methods

Study Organisms

Water striders of the genus *Limnoporus* Stål (Heteroptera: Gerridae) inhabit standing waters throughout the northern part of the Holarctic region. Andersen and Spence (1992) recognized six species in their taxonomic revision of the genus and conducted a cladistic analysis based on 45 structural characters, using species of *Gerris* and *Aquarius* as outgroups. The phylogenetic analysis revealed two monophyletic subgroups within the genus (Fig. 4-1): the *L. canaliculatus* group, consisting of the two small species *L. canaliculatus* and *L. esakii* (total body length less than 11.5 mm), and the *L. rufoscutellatus* group with the remaining four species, which are all clearly larger. Allozyme analysis (Sperling and Spence, 1990) also substantiates the close relationships among *L. notabilis*, *L. dissortis*, and *L. rufoscutellatus*, and their isolation from *L. canaliculatus*. The divergence of the *L. canaliculatus* and *L. rufoscutellatus* groups probably occured before about 50 MYA, as evidenced by fossils from the middle Eocene (Andersen et al., 1993). The *L. canaliculatus* group has a disjunct distribution: *L. canaliculatus* occurs in eastern North America, and *L. easkii* in East Asia (Andersen and Spence, 1992). This suggests that the two species diverged before the late Miocene drop in global temperature (Briggs, 1987; Zubakov and Borzenkova, 1990).

Like most semiaquatic bugs, all *Limnoporus* species have five larval instars, which can be considered as homologous ontogenetic stages (Andersen, 1982). The cuticle of the legs and antennae is rigidly sclerotized in all larval instars, and unable to expand between molts. Growth of these structures therefore proceeds in a stepwise manner, and makes them especially suitable for the quantitative study of ontogeny.

Data

Morphometric measurements were made on all five larval instars (denoted L1 to L5, respectively) and adults of all six *Limnoporus* species reared in mass cultures in the laboratory. Cultures were maintained under long-day photoperiods at 23–25°C, and fed frozen flesh flies (*Sarcophaga bullata* Parker) six times per week. The origins of the laboratory cultures are as follows: *L. notabilis* from Vancouver, British Columbia, Canada; *L. dissortis* from Edmonton,

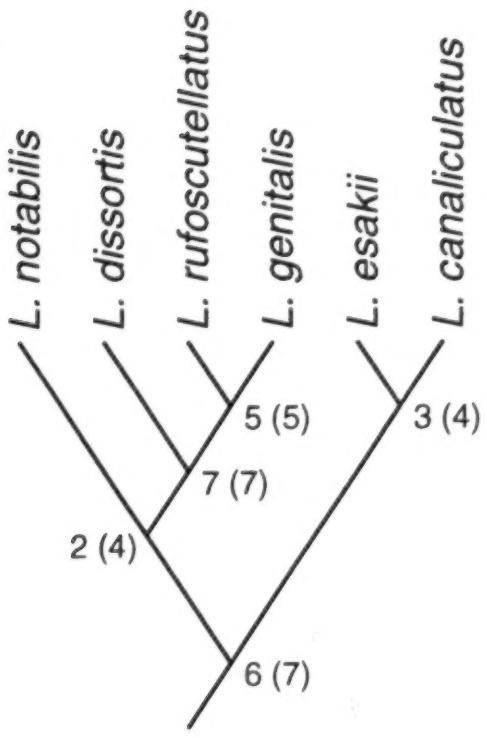


Figure 4-1. Hypothesized phylogeny of the genus *Limnoporus* according to Andersen and Spence (1992). The number of synapomorphies unrelated to the traits used in the present study (lengths of antennal and leg segments) is indicated for each node (total number of synapomorphies in parentheses).

Alberta, Canada; *L. rufoscutellatus* from Hanko near Tvärminne, Finland; *L. genitalis* from Hokkaido, Japan; *L. esakii* from Honshu, Japan; *L. canaliculatus* from Kingston, Ontario, Canada.

For each species, 10 specimens were measured for each of the first three instars, where the sexes could not be identified, and 10 specimens per instar and sex for the L4, L5 and adult (for *L. esakii*, 7 L2 and 11 L3 were measured). Therefore, the morphometric data are of true cross-sectional type (Cock, 1966). Specimens were preserved in alcohol before measuring. Measurements were made with a dissecting microscope equipped with an eyepiece micrometer. Eight measurements are included in this study: the lengths of all four antennal segments (ANTSEG1 to ANTSEG4) and the lengths of femora and tibiae of the middle and hind legs (MIDFEM, MIDTIB, HINDFEM, and HINDTIB, respectively).

As a measure of chronological time, instar durations were determined for larvae that were reared individually. All six species were reared in the same room under long-day photoperiod (19L : 5D) at 20°C. Each bug was fed a flesh fly

daily, and checked for molts at intervals of approximately 12 hours. Due to high mortality, specimens taken from the mass culture as L4 were also used to estimate the mean duration of the L5 for *L. dissortis*. Eggs or larvae were taken from the same laboratory cultures as for morphometric measurements, except for *L. canaliculatus*, for which instar durations were determined using a sample from a population from Morris County, New Jersey, USA.

Statistical Analysis

Multivariate Allometry.—Jolicoeur (1963) proposed the first principal component of the covariance matrix of log-transformed measurements as a multivariate generalization of the allometry equation. Applied to measurements made on individuals of a single species differing in age, the first principal component approximates the ontogenetic trajectory in the space defined by the morphometric variables (Shea, 1985; Klingenberg and Zimmermann, 1992b). The coefficients of the first principal component reflect the direction of the ontogenetic trajectory, and are roughly proportional to the slopes obtained in bivariate allometric regressions of the traits on a measure of overall size (Davies and Brown, 1972; Shea, 1985). Therefore, they can be interpreted as patterns of multivariate allometry. We computed principal components for both sexes of each species separately, using the covariance matrix of log-transformed data pooled over ontogenetic stages. Because the sexes could not be distinguished in the first three instars, the samples of L1–L3 were included in the analyses for both sexes.

The bootstrap technique (Efron and Tibshirani, 1986) was used to assess statistical accuracy of principal component estimates. Bootstrap samples were taken separately for each stage before pooling these samples and computing principal components. For each species and sex, 2,000 bootstrap iterations were performed.

As a graphical comparison of the directions of ontogenetic trajectories, we used an ordination of the allometric patterns of all the species and both sexes (Klingenberg and Froese, 1991). The coefficients of the first principal components within the 12 groups were used as "observations" in another principal component analysis. The scores of the groups on first and second of the resulting component axes were plotted to display the variation among species and sexes in the directions of their ontogenetic trajectories.

Common principal components (Flury, 1988; Airoldi and Flury, 1988; Klingenberg and Zimmermann, 1992b) were used as a joint estimate of the direction of growth trajectories for all 12 groups. For estimating common principal components, we used a version of the FG-algorithm (Flury, 1988) written in SAS/IML language (a listing is available from C.P.K. on request). Statistical accuracy was assessed with a bootstrap approach corresponding to the one used for one-group principal component analyses.

Lateral Transpositions of Growth Trajectories.—To compare the relative roles of ontogenetic scaling and lateral transpositions of trajectories, and to assess the effects of non-allometric growth, we separated morphometric variation along growth trajectories from variation in transverse directions using Burnaby's technique of adjusting for growth (Burnaby, 1966; Rohlf and Bookstein, 1987). This technique removes variation in the direction of a growth vector **b** representing the ontogenetic trajectories, and is equivalent to setting within-group size to a constant value (e.g., zero) and considering only the

variation in directions perpendicular to **b**. For a data matrix **X** ($n \times p$) consisting of n observations and p variables, adjusted data are calculated by postmultiplying **X** by the matrix $\mathbf{Q} = \mathbf{I}_p - \mathbf{b}(\mathbf{b}'\mathbf{b})^{-1}\mathbf{b}'$, where $\mathbf{I}_p$ is an identity matrix of rank p (for a normalized vector, such as a principal component, the expression simplifies to $\mathbf{Q} = \mathbf{I}_p - \mathbf{b}\mathbf{b}'$). Instead of the first principal component of the pooled within-group covariance matrix, as recommended by Rohlf and Bookstein (1987), we used the first common principal component as an estimate for the direction of growth trajectories (see Airoldi and Flury [1988] for discussion).

Besides lateral shifts of the entire trajectories, non-allometric growth may also contribute to morphometric variation orthogonal to the ontogenetic trajectories, e.g., the slight curvatures of ontogenetic trajectories observed in gerrids (Klingenberg and Zimmermann, 1992b, and references therein). As an attempt to separate these two sources of variation, we used the data adjusted with Burnaby's technique in a two-way MANOVA with species and instar as "treatment" factors. The first eigenvectors (principal components) of the matrices of sums and cross-products due to the two main effects accounted for most of the variation in both matrices. Therefore, they were used to display graphically the variation due to lateral shifts of trajectories (between-species effects) and non-allometric growth (between-instar effects). Because data adjusted for growth by Burnaby's technique have singular covariance matrices and cannot be used for statistical tests if the growth vector **b** is derived from the same data (Burnaby, 1966), we used the MANOVA and resulting principal components only as an ordination.

Size Increments.—Whereas multivariate allometry pertains to the direction of ontogenetic trajectories in the space defined by morphometric variables, the relative positions of life history stages along the trajectory are another important aspect of growth, because they can be interpreted as a multivariate measure of size. We defined a measure of overall size by rescaling the first common principal component so that its coefficients sum up to unity. This size measure scales as a linear dimension, and therefore fulfills Mosimann's (1970) definition of standard size variables (for discussion, see Klingenberg and Zimmermann [1992a]).

Size increments were calculated as the geometric-mean growth ratio for each molt, i.e., the ratio of the geometric means of the multivariate size measure in two successive instars. This ratio, which can be obtained from cross-sectional data, is equivalent to the geometric mean of the ratios of the size measure calculated for individual bugs in the two instars, as they would be calculated from longitudinal data (Klingenberg and Zimmermann, 1992a). Confidence intervals of geometric-mean growth ratios were established using a bootstrap procedure with 2,000 iterations for each molt (for details, see Klingenberg and Zimmermann [1992a]).

Total development time (only for specimens with complete data from hatching to the imaginal molt) was compared among species and sexes with the Tukey-Kramer procedure for multiple comparisons of means (Sokal and Rohlf, 1981).

Table 4-1. Ontogenetic allometry of both sexes of the six *Limnoporus* species. Tabled values are coefficients of the first principal components of log-transformed measurements and bootstrapped standard errors (in parentheses).

	L. notabilis		*L. dissortis*		*L. rufoscutellatus*		*L. genitalis*		*L. esakii*		*L. canaliculatus*	
Variable	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males
ANTSEG1	0.418	0.430	0.392	0.395	0.385	0.404	0.397	0.403	0.406	0.404	0.390	0.399
	(0.004)	(0.004)	(0.002)	(0.003)	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
ANTSEG2	0.395	0.393	0.386	0.391	0.383	0.380	0.382	0.396	0.398	0.394	0.387	0.397
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
ANTSEG3	0.339	0.329	0.318	0.319	0.329	0.336	0.314	0.322	0.313	0.310	0.331	0.326
	(0.004)	(0.007)	(0.003)	(0.003)	(0.009)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
ANTSEG4	0.206	0.200	0.188	0.183	0.221	0.223	0.192	0.194	0.179	0.167	0.188	0.188
	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)
MIDFEM	0.372	0.378	0.372	0.367	0.380	0.379	0.367	0.362	0.372	0.370	0.369	0.363
	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
MIDTIB	0.309	0.305	0.311	0.309	0.296	0.289	0.305	0.297	0.301	0.313	0.307	0.301
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)
HINDFEM	0.396	0.404	0.403	0.404	0.397	0.394	0.406	0.404	0.394	0.402	0.400	0.399
	(0.007)	(0.008)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
HINDTIB	0.349	0.338	0.405	0.403	0.399	0.383	0.410	0.397	0.403	0.402	0.403	0.400
	(0.007)	(0.008)	(0.002)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)

RESULTS

Directions of Ontogenetic Trajectories

The estimates of first principal component coefficients (Table 4-1) are fairly stable, as can be seen from their small standard errors, and most of them clearly differ from 0.354, the value for isometry. There are some marked differences among species, especially for ANTSEG4 and HINDTIB. Overall, however, differences are small, as indicated by the narrow angles between trajectories of different groups (maximum angle: 4.72° between *L. canaliculatus* males and *L. genitalis* females, corresponding to a component correlation of 0.997). The proportion of total variance taken up by the first principal component varies from 98.0% (*L. esakii* females) to 99.4% (*L. notabilis* males and both sexes of *L. rufoscutellatus*). Therefore, most ontogenetic variation is contained in a single dimension, and is described fairly accurately by the patterns of ontogenetic allometry as given by the first principal components.

The ordination of allometric patterns (Fig. 4-2) is remarkably consistent with the hypothesized phylogeny of the genus (Fig. 4-1). The two sister species *L. canaliculatus* and *L. esakii* are well separated from one another and from the

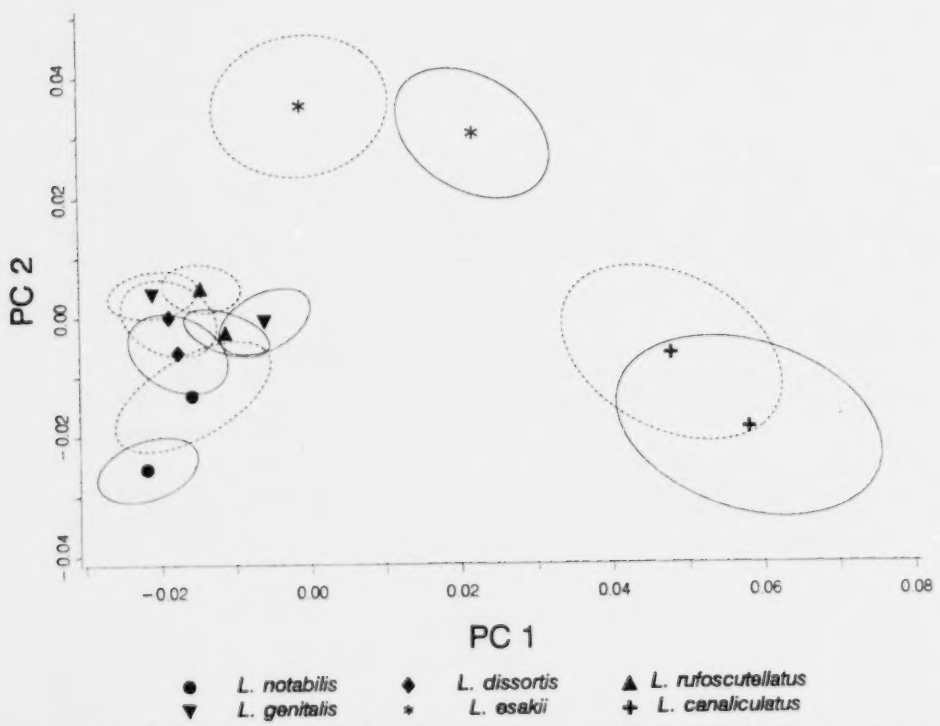


Figure 4-2. Ordination of multivariate allometric patterns by principal components. Principal component scores of the estimates of allometric patterns for each species and sex are graphed with 95% confidence ellipses derived from the respective bootstrap analyses. Dashed lines: females. Solid lines: males.

Table 4-2. Estimates of the first common principal component (CPC1) and bootstrapped standard errors.

Variable	CPC1	S.E.
ANTSEG1	0.398	0.0010
ANTSEG2	0.391	0.0010
ANTSEG3	0.324	0.0013
ANTSEG4	0.193	0.0012
MIDFEM	0.370	0.0008
MIDTIB	0.303	0.0008
HINDFEM	0.401	0.0007
HINDTIB	0.396	0.0013

dense cluster formed by the species of the *L. rufoscutellatus* group. Within that cluster, *L. notabilis*, and to a lesser extent also *L. dissortis*, are somewhat set off from *L. rufoscutellatus* and *L. genitalis*.

Estimated coefficients of the first common principal component (Table 4-2) are similar to the first principal components of separate groups. Angles between one-group principal components and the common principal component vary from 0.67° (*L. rufoscutellatus* males) to 3.88° (*L. canaliculatus* males). Between 97.8% (*L. esakii* females) and 99.4% (*L. rufoscutellatus* males) of the total variation within each group are taken up by the first common principal component; this is almost as much as in the analyses of individual groups. The first common principal component can therefore be considered as a good joint estimate of the direction of growth trajectories.

Ontogenetic Scaling and Lateral Transposition

Data from which we had removed variation in the common direction of growth trajectories using Burnaby's procedure still contained a mixture of variation due to lateral shifts of growth trajectories and to non-allometric growth. The separation of these effects with a two-way MANOVA produced between-species and between-instar matrices whose first principal components were almost orthogonal (angle 92.8° and component correlation –0.048). Nevertheless, the procedure did not completely separate the two sources of variation because of species × instar interaction suggested by Figures 4-3 and 4-4.

Effects of lateral transposition are displayed in a plot of the between-species component against the common growth axis (Fig. 4-3). The trajectories of *L. canaliculatus* and *L. esakii* are separated from those of the other four species by an upward lateral transposition, but also by a shift to the left, which indicates ontogenetic scaling toward smaller sizes at all growth stages. Lateral transpositions also occur among the four species of the *L. rufoscutellatus* group, but the picture is complicated by non-allometric growth and by sexual dimorphism (especially conspicuous in L5 of *L. genitalis*).

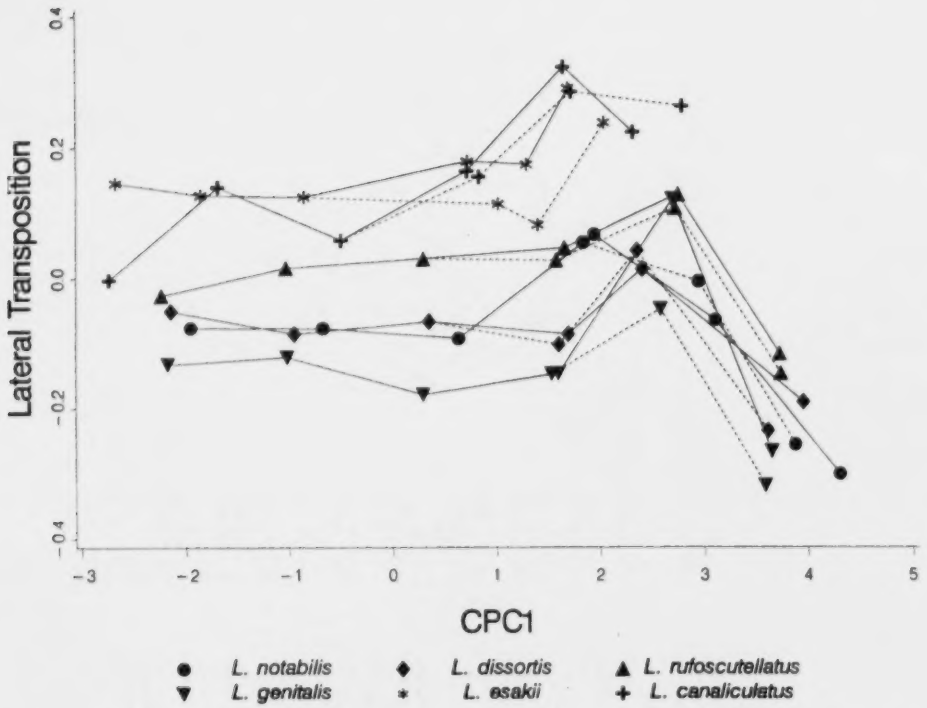


Figure 4-3. Morphometric variation due to lateral transposition of ontogenetic trajectories. The axis labeled "Lateral Transposition" is the first principal component of the between-species matrix of sums of squares and cross-products in a two-way MANOVA (species × instars) of data from which variation along the growth trajectories has been removed by Burnaby's procedure. The first common principal component (CPC1) is a joint estimate of the direction of growth trajectories, with L1 on the left and adults to the right side. Plotted points are mean scores for each species, instar, and sex. Dashed lines: females. Solid lines: males. Note the difference in scale between the two axes.

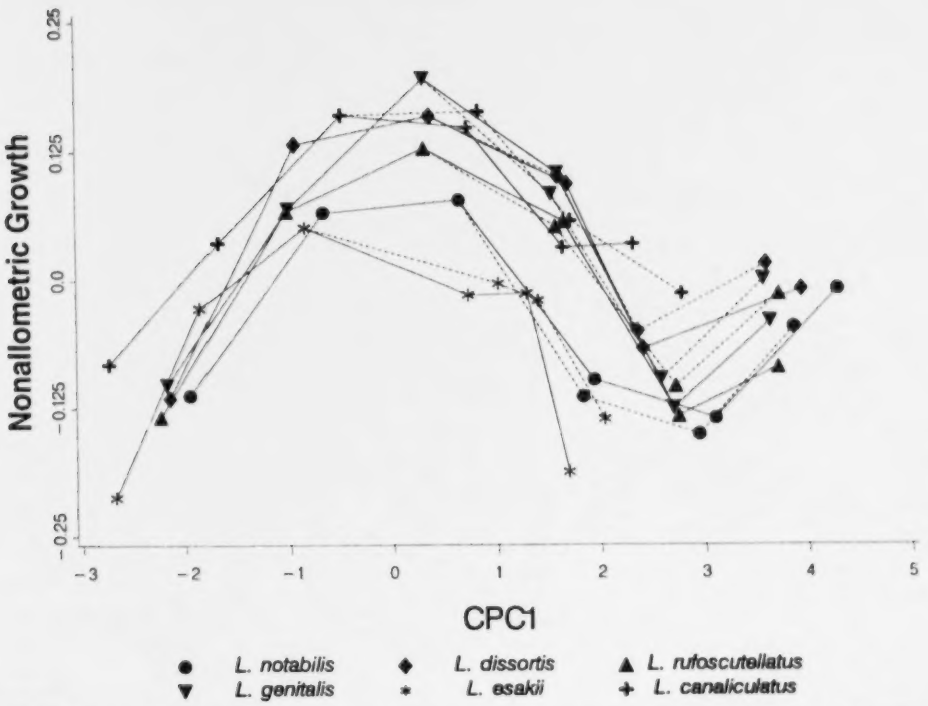


Figure 4-4. Morphometric variation due to non-allometric growth. The axis labeled "Nonallometric Growth" is the first principal component of the between-instars matrix of sums of squares and cross-products in a two-way MANOVA (species × instars) of growth-adjusted data. The first common principal component (CPC1) is a joint estimate of the direction of growth trajectories. Plotted points are mean scores for each species, instar, and sex. Dashed lines: females. Solid lines: males. Note the difference in scale between the two axes.

A component of non-allometric growth exists in all six species, as can be seen from the consistently upward-convex curvature of trajectories in Figure 4-4. In addition, trajectories of all four species of the *L. rufoscutellatus* group turn upward between the L5 and the adult stage (Fig. 4-4), and, at the same time, decrease sharply in their scores for lateral transposition (Fig. 4-3).

Size Increments

Geometric-mean growth ratios vary within and between species (Fig. 4-5). Especially *L. esakii*, *L. dissortis*, and to a lesser degree also *L. canaliculatus*, show marked differences in growth increments between molts. In the other species, growth ratios vary within a limited range only. In most species where sexual dimorphism occurs, it is most conspicuous in the last molt (*L. notabilis*, *L. dissortis*, *L. canaliculatus*). In *L. esakii*, however, growth increments at the molt to L4 differ more between sexes than at the molt to L5. There is no apparent correspondence between these differences in patterns of size increments and the phylogenetic relationships of the six species (Fig. 4-1).

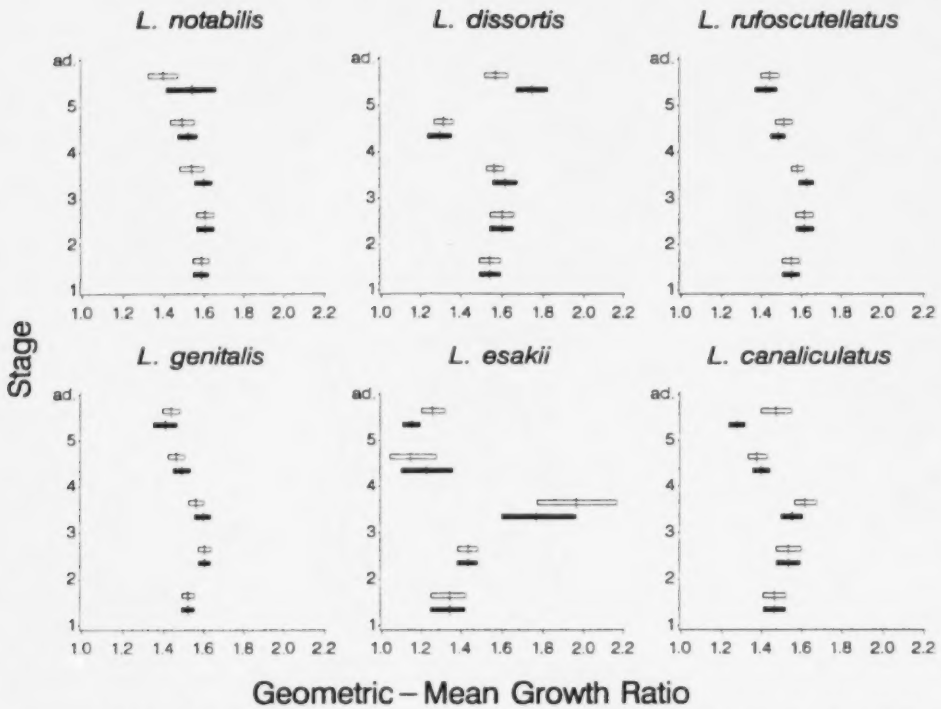


Figure 4-5. Size increments. Plotted values are geometric-mean growth ratios for the multivariate size measure and their bootstrapped central 95% confidence intervals (bars). Open bars: females. Solid bars: males.

Heterochrony

To test the association between size and age, we plotted geometric-mean growth ratios based on the multivariate size measure against instar duration (Fig. 4-6). There is no apparent overall relationship between the two variables ($r = -0.17$, $N = 60$, $P > 0.19$, two-tailed). Within all species except *L. dissortis*, correlations are negative, as later instars tend to have relatively small growth increments (Fig. 4-5) and to last longer than earlier instars (Table 4-3).

Because homologous developmental events (the molts in our example) as well as external time can be used as a reference for the study of heterochrony, we considered both these measures. Graphs of the log-transformed multivariate size measure against instar number (Fig. 4-7a) are nearly linear (except for *L. esakii* and *L. dissortis*) and have similar slopes. This reflects the limited variation of growth ratios (Fig. 4-5).

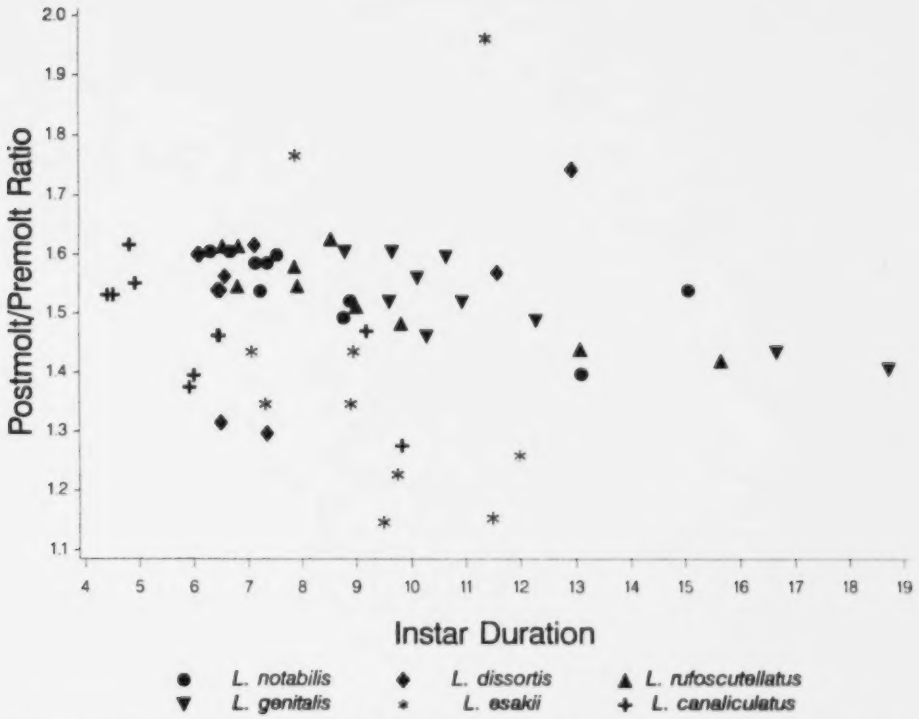


Figure 4-6. Size increments graphed against instar durations. Postmolt/premolt ratios are geometric-mean growth ratios based on the multivariate measure of size. Sexes are graphed separately for each species.

The graphs of size against age (Fig. 4-7b) are all slightly curved, indicating changes in growth rate, and there are marked differences among species. Within the *L. rufoscutellatus* group, the growth curves are similar, except for *L. genitalis*, which has a longer developmental time and lower growth rates than the other three species (Tukey-Kramer test: males differ significantly [$P < 0.05$] from all other *Limnoporus* species, females from all other species except *L. esakii* females). Using the principle of parsimony, we hypothesize that the common ancestor of *L. genitalis* and *L. rufoscutellatus* had a growth curve similar to *L. rufoscutellatus*, *L. dissortis*, and *L. notabilis* (cf. Fig. 4-1). Therefore, *L. genitalis* is both hypermorphic and neotenous relative to its hypothetical ancestor. *L. canaliculatus* has an extremely short development time (Tukey-Kramer test: both sexes differ significantly from all other species except *L. dissortis* females) and has also higher growth rates (especially during L2 and L3; Fig. 4-7b). Progenesis and acceleration were the major heterochronic processes in the evolution of *L. canaliculatus*. It is not clear at which place in the phylogeny acceleration occurred, because *L. esakii* also shows high growth rates during the L3, and growth rates of the common ancestor of the entire genus cannot be inferred. *L. esakii* has extremely low growth rates in the last two instars, indicating neoteny.

Table 4-3. Mean instar durations of the six *Limnoporus* species. Tabled values are means, standard errors and sample sizes (in parentheses) are given in the second line of each entry.

Species	Sex	Instar				
		L1	L2	L3	L4	L5
L. notabilis	f	7.1	6.3	7.2	8.8	13.1
		0.34 (11)	0.30 (11)	0.52 (11)	0.57 (11)	0.33 (9)
	m	7.4	6.7	7.5	8.9	15.1
		0.18 (14)	0.14 (14)	0.23 (14)	0.44 (14)	0.69 (10)
L. dissortis	f	6.5	6.1	6.6	6.5	11.6
		0.00 (6)	0.51 (6)	0.47 (6)	0.22 (6)	0.33 (5)
	m	6.4	6.1	7.1	7.3	13.0
		0.12 (16)	0.27 (16)	0.38 (16)	0.35 (16)	0.65 (12)
L. rufo-scutellatus	f	7.9	6.8	7.9	9.0	13.1
		0.31 (12)	0.21 (12)	0.29 (11)	0.29 (11)	0.48 (11)
	m	6.8	6.5	8.5	9.8	15.7
		0.17 (13)	0.24 (13)	0.43 (13)	0.64 (13)	1.06 (8)
L. genitalis	f	9.6	8.8	10.1	10.3	16.7
		0.61 (10)	0.63 (10)	0.35 (9)	0.35 (9)	0.62 (10)
	m	11.0	9.7	10.7	12.3	18.8
		1.23 (10)	0.65 (10)	0.45 (10)	0.70 (10)	0.73 (9)
L. esakii	f	8.9	9.0	11.4	9.5	12.0
		0.49 (10)	0.99 (10)	1.51 (10)	0.35 (10)	0.55 (10)
	m	7.3	7.1	7.9	9.8	11.5
		0.28 (8)	0.79 (8)	1.01 (8)	1.04 (8)	0.46 (8)
L. canali-culatus	f	6.5	4.4	4.8	5.9	9.2
		0.04	0.06	0.06	0.07	0.11
		(150)	(150)	(150)	(150)	(119)
	m	6.5	4.5	4.9	6.0	9.8
		0.04	0.06	0.07	0.07	0.15 (97)
		(153)	(153)	(153)	(151)	

The plot of instar number against age (Fig. 4-7c) again shows the almost twofold difference in development time between the extreme species *L. canaliculatus* and *L. genitalis*. There are marked differences between sexes as well. The clearest example is for *L. esakii*, in which females tend to have considerably longer development times (Tukey-Kramer test, $P < 0.05$) and slightly lower growth rates (Fig. 4-7b) than males. In this species, as well as in *L. notabilis* and *L. dissortis*, sexual size differences seem to be produced mostly by hypermorphosis of the larger sex (females in *L. esakii*, males in *L. notabilis* and *L. dissortis*) relative to the smaller sex in at least one instar (L3 in *L. esakii*, L5 in *L. notabilis* and *L. dissortis*). In *L. canaliculatus*, however, the lower growth rate of males in the L5 produces the size difference between sexes (the difference in development time is not significant despite the large

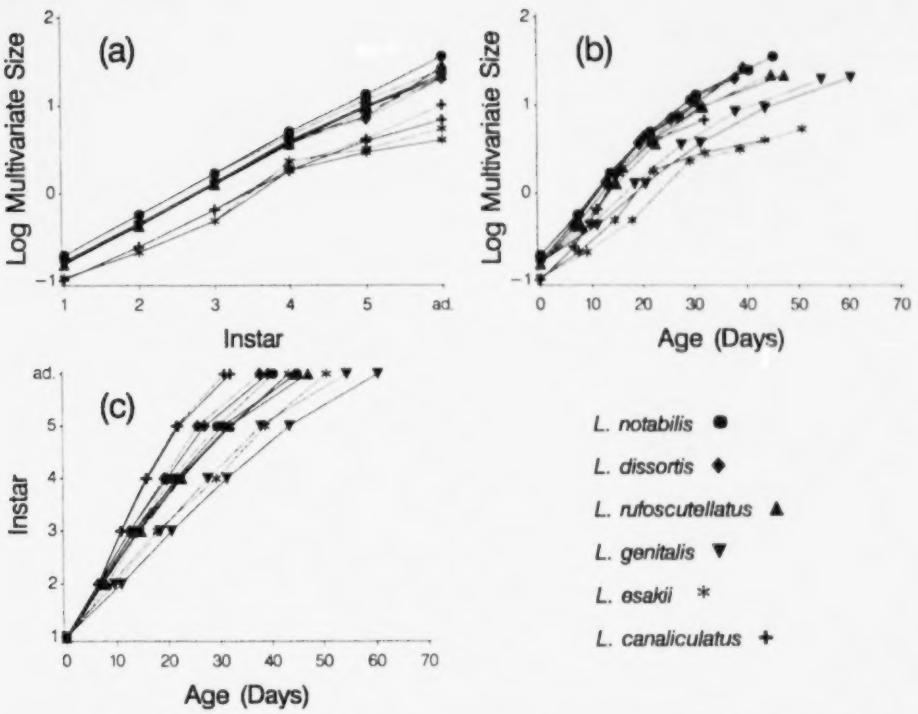


Figure 4-7. Relationships between the multivariate size measure, age, and instars. (a) The log-transformed size measure plotted against instar number. (b) The log-transformed size measure graphed against age (days after hatching) at the beginning of the corresponding ontogenetic stage. (c) Plot of instar number against age.

sample size). Conversely, in *L. genitalis*, development time differs markedly between sexes (although not statistically significant), but there is very litte difference in size (Fig. 4-7a, b).

DISCUSSION

Heterochrony explains evolutionary changes in form by changes in the rate or timing of the developmental processes that produce the structures of interest. Allometry, on the other hand, characterizes patterns of character variation and associations among traits; changes in allometries may be the result of heterochronic alterations. As demonstrated by our simple model (Fig. 3-6), allometric patterns cannot be used to infer heterochronic processes. Rather, allometry and heterochrony are conceptually distinct, complementary parts in a comprehensive analysis of the evolution of form.

We used multivariate allometry and the time intervals between discrete developmental events (molts) to assess patterns of character variation and the role of heterochrony in the evolution of a clade of water striders. Our analyses revealed variation in the directions of growth trajectories, lateral transposition of trajectories, ontogenetic scaling, and some major heterochronic changes; moreover, all

these effects are tightly interwoven with non-allometric growth and sexual dimorphism.

Allometry

The first principal components account for almost the entire ontogenetic variation; growth trajectories are almost straight lines in the space of log-transformed measurements. Despite the close similarity indicated by the narrow angles between ontogenetic trajectories, allometric patterns differ significantly among the six species (Fig. 4-2). These differences correspond remarkably well to the hypothesized phylogeny of the genus (Fig. 4-1) as proposed by Andersen and Spence (1992). This correspondence can be interpreted as an indication of divergent evolution of the traits included in this study. There is, however, a possible alternative interpretation: the correspondence may be artifactual because some of the characters used by Andersen and Spence (e.g., "fourth antennal segment shorter than first segment") are based on the same traits as are included in the present study. This alternative can be ruled out because all the nodes of the cladogram are supported by at least two qualitative characters that are unrelated to the mensural traits used here (Fig. 4-1), and presumably are independent of allometric trends or size-scaling. Independent evidence from an allozyme study (Sperling and Spence, 1990) also supports the topology of the cladogram. Consequently, the phylogenetic hypothesis of Andersen and Spence (1992) is a robust base of comparison. Interspecific variation in the directions of growth trajectories, reflecting changes in the relative growth rates of the traits measured (Jolicœur, 1963; Shea, 1985), is thus an indication of morphological divergence among *Limnoporus* species.

Using the first principal component to characterize allometric patterns is equivalent to fitting a straight line to the growth data, and therefore yields an "overall direction" of the ontogenetic trajectories, disregarding nonlinearities caused by fluctuations in relative growth rates during ontogeny (non-allometric growth). Although they only accounted for a minor fraction of the total morphometric variation, two main features of non-allometric growth emerged: a curvature from the L1 to L5 instars in all species, and a sharp twist between the L5 and the adult stage (Fig. 4-4). The discrepancy in allometric patterns between the *L. canaliculatus* and *L. rufoscutellatus* groups (Fig. 4-2) may be due in part to the twist between the L5 and adult stages, which is present in all four species of the *L. rufoscutellatus* group, but very weak in *L. canaliculatus* and not detectable in *L. esakii* (Figs. 4-3, 4-4).

Curvatures of growth trajectories in the space defined by log-transformed measurements were found in earlier multivariate studies of growth in other gerrids (Klingenberg and Zimmermann, 1992b) and in a backswimmer (Cuzin-Roudy and Laval, 1975). Cuzin-Roudy and Laval (1975) and, using untransformed data, Blackith et al. (1963) and Davies and Brown (1972), found that this curvature was more accentuated in the last instar, but in the same direction as in earlier instars. In contrast, our data for *Limnoporus* show that the last molt produces a twist in the direction opposite to the curvature in the earlier stages in the four larger *Limnoporus* species (Fig. 4-4). A similar bend in allometric trajectories has only been described for a supernumerary larval instar of a backswimmer produced by treatment with a juvenile hormone analogue, where

trajectories turned in a direction opposite to the curvature in earlier instars (Cuzin-Roudy and Laval, 1975).

The statistical technique we used to display lateral transposition of ontogenetic trajectories and non-allometric growth is specifically designed to separate variation along trajectories from variation orthogonal to them (Burnaby, 1966; Rohlf and Bookstein, 1987). Shea (1985) proposed to use a principal component analysis on pooled samples for this purpose. If lateral transposition co-occurs with a shift along the trajectories in some groups, the first principal component of pooled samples may intersect trajectories at oblique angles (Fig. 3-30). Shea (1985) argued that the oblique orientation of trajectories might be due to differences in their directions among groups. As Figure 4-8 shows, however, this is not necessarily true; Shea's explanation does not apply if trajectories are parallel, but intersect the "total" first principal component at an oblique angle. This is because the "total" first principal component is the direction accounting for the most variation, regardless of whether the variation is within or between groups. Because lateral transposition is orthogonal to the trajectories by definition, the statistical technique used to analyze it should reflect this situation, as Burnaby's procedure does.

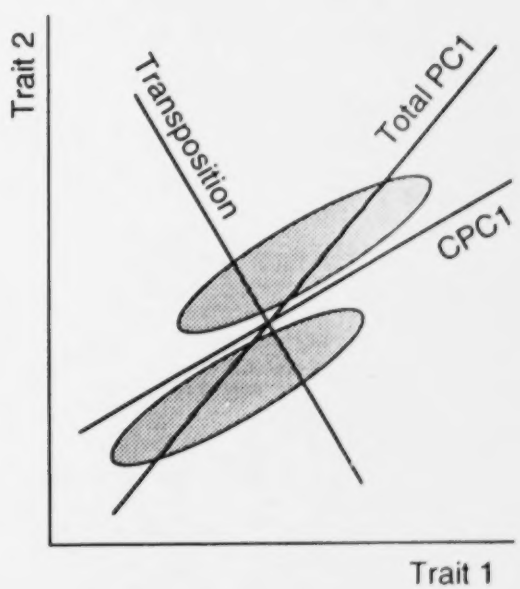


Figure 4-8. Principal components and lateral transposition of growth trajectories. If growth trajectories are parallel, the first common principal component (CPC1) indicates their direction. An axis of lateral transposition can be found as the direction of maximal variation between groups, subject to the constraint that it is orthogonal to the CPC1 (this is achieved by Burnaby's technique). If differences between groups are based on lateral transposition and ontogenetic scaling (shifts along the growth trajectories), then the first principal component of the pooled samples (Total PC1) intersects the growth trajectories at an oblique angle and confounds ontogenetic variation within groups with differences between groups.

A clear lateral transposition of allometric trajectories and a shift of all stages along the trajectories toward smaller sizes separate the *L. canaliculatus* group from the *L. rufoscutellatus* group (Fig. 4-3). Smaller changes of this kind can also be seen within the *L. rufoscutellatus* group. While these interspecific differences can be described as shifts of allometric growth trajectories, it is not possible to identify the heterochronic processes responsible. Pre- or post-displacement can lead to lateral transpositions of growth trajectories (Fig. 3-6); in our study, however, all traits already have started growth in the first stage considered (L1). Therefore, the heterochronic processes that produced these transpositions of trajectories must have acted before hatching; embryological data would be necessary to identify the nature of these processes. The same applies to the shifts along growth trajectories, although global heterochronies affecting any later stages may also contribute to ontogenetic scaling by extending or truncating common allometries (Fig. 3-6).

Heterochrony

The relations between age and size revealed by our case study are complex, rather than a simple linear dependence (Figs. 4-6, 4-7). The correlation between age and size is bound to be positive because both age and size are monotonically increasing, but this correlation need not be biologically meaningful. A rigorous test of the relation of size and age must focus on the increments in size and time during stages (Fig. 4-6), instead of the cumulative curves (Fig. 4-7b). There is no positive correlation between growth increments and instar durations, neither within nor between species (Fig. 4-6). Therefore, size cannot be taken as a proxy for age. For the same reason, however, heterochronic changes of size itself become a focus for study.

No major heterochronic changes apparent in our data (Fig. 4-7b) correspond to the "pure" processes (Fig. 3-3) of the theoretical scheme by Alberch et al. (1979). Rather, combinations of two or more of the processes act simultaneously, partially compensating for each other's effects (e.g., neoteny and hypermorphosis in *L. genitalis*, acceleration and progenesis in *L. canaliculatus*). Processes that affect only part of the ontogeny (e.g., the neoteny of *L. esakii* in the last two instars) also add further complexity to the analysis of heterochrony in *Limnoporus*.

An important problem in a study of heterochrony is the choice of a metric for time (e.g., Schmidt-Nielsen, 1984; Blackstone and Yund, 1989; Reiss, 1989; McKinney and McNamara, 1991). The question is whether internal (physiological) time, external (clock) time, or a measure based on homologous developmental events (molts in our example) is most appropriate (e.g., Blackstone, 1987b; Reiss,1989). Internal time of ectothermic animals depends on ambient temperature (Spence et al., 1980; Taylor, 1981). Because our data on development time were obtained under standardized laboratory conditions, temperature does not contribute to the variation in our data. Small animals generally tend to develop more quickly than larger species (Schmidt-Nielsen, 1984; Reiss, 1989). This is not the case in our example: *L. genitalis* is not larger than the other species of the *L. rufoscutellatus* group, yet has a clearly longer development time (Fig. 4-7b, c). The deviation from the general rule is even more striking in *L. esakii* and *L. canaliculatus*, where the smaller *L. esakii* has a much longer development time than its somewhat larger-bodied sister species *L.*

canaliculatus. Furthermore, if the durations of different instars are compared to study the allocation of time to homologous developmental stages (Fig. 4-7c), there is also considerable variation among species, apparently unrelated to their phylogeny. Therefore, the variation in development time among the six *Limnoporus* species is not a consequence of their variation in size or of phylogenetic inertia, but possibly reflects adaptive evolution of this life history trait. Because the corresponding population processes, such as the rate of mortality by predation, are measured on an extrinsic time scale, we use clock time as a reference dimension for heterochrony (see also Blackstone, 1987b).

The possible adaptive causes for the prolonged developmental time in *L. genitalis* are unclear. We can conceive an adaptive scenario, however, that would explain the heterochronic changes in *L. canaliculatus* and *L. esakii*. Selection for shorter development time by high larval mortality may be the cause of the combination of acceleration and progenesis observed in *L. canaliculatus*. This heterochronic innovation appeared after the speciation event from which *L. canaliculatus* and *L. esakii* originated (Fig. 4-1), and was therefore not available to the latter species. The sexual dimorphism of *L. esakii* evolved as a response to a trade-off between selection for rapid completion of development, even at small size (in males), and selection for larger size in females because of the association between female size and fecundity (corresponding to the "developmental constraints" hypothesis of Fairbairn [1990]). This scenario, however, does not account for the small size increments in the final two instars of *L. esakii* (Fig. 4-5). Because it refers to evolutionary events in the past, which occurred under climatic conditions different from the present (Zubakov and Borzenkova, 1990), the hypothesis as a whole is not testable. A partial test in the field might focus on the maintenance of sexual dimorphism in *L. esakii*: the hypothesis predicts that males have a higher larval survivorship than females, and that female size correlates positively with lifetime fecundity. Although no direct evidence is available for these two species, two field studies of similar-sized gerrids document high larval mortalities (Zimmermann et al., 1982; Spence, 1986). Fairbairn (1988) found low, but significant correlations between female body length and the size of egg batches in three water strider species; however, no reliable evidence about the relation between female size and lifetime fecundity has been published for water striders (Spence and Andersen, 1994).

Growth dynamics (Fig. 4-7) and allocation of growth to different instars (Fig. 4-5) differ greatly between species. In a similar study in nine water strider species of the genera *Gerris* and *Aquarius*, geometric-mean growth ratios of a multivariate size measure (including three additional characters) varied only between 1.38 and 1.58 (Klingenberg and Zimmermann, 1992a), which is a considerably narrower range than we observed here (Fig. 4-5). The variability within the genus *Limnoporus* indicates a fair amount of evolutionary plasticity of developmental processes and the associated life-history traits. The role of genetic constraints for the evolution of ontogenetic trajectories (e.g., Cheverud et al., 1983; Kirkpatrick and Lofsvold, 1992) is unclear in the absence of genetic data. Whereas the variability observed among species suggests a considerable evolutionary potential, the scenario outlined above emphasizes the possibility that in *L. canaliculatus* an innovation may have overcome constraints that still exist in *L. esakii*.

In most species, the sexes differ in growth increments (Fig. 4-5) and development times (Fig. 4-7c). These differences are not necessarily linked to sexual size dimorphism in adults (e.g., *L. genitalis*). Moreover, in the species where size dimorphism exists, it is achieved by changing growth rates or durations in different ontogenetic stages. This result is consistent with the failure of unitary hypotheses to explain sexual size dimorphism in gerrids (Fairbairn, 1990). Specific adaptations of gerrid life histories to environmental conditions (e.g., Spence, 1989) seem to predominate over general constraints. In the phylogeny of *Limnoporus*, over a time scale of millions of years, various ways to dissociate the ontogenetic trajectories of the two sexes have been available.

Given this variety of patterns, even within a small clade, it is necessary to carry out a detailed analysis for each specific case in order to identify the processes involved in evolutionary changes. "Global" tests of hypotheses across a spectrum of species are likely to fail, whether or not they are actually applicable in specific instances. An approach that seems more promising involves a synthesis of detailed ecological information and studies of growth and form on the background of a well-resolved phylogenetic hypothesis.

Conclusions

To understand the connection between heterochrony and allometry, it is necessary to distinguish clearly between patterns and processes. Allometry—the variation and covariation of characters in the space spanned by measurements of form—is a characterization of a *pattern*, which is the result of the underlying developmental phenomena. The model of heterochrony proposed by Alberch et al. (1979), on the other hand, is based on a model of a developmental *process*, and simple changes of the model parameters. Therefore, it can be used as a formalism to accurately describe evolutionary changes in ontogenetic pathways, which may help us understand the causes and consequences of those changes. It is less useful, however, as a classification scheme, because the heterochronic processes in this framework are not mutually exclusive (except those affecting the same parameter in opposite directions). Our case study of heterochrony in *Limnoporus*, where none of the major heterochronic changes corresponds to a "pure" process, illustrates the importance of the combined action of several processes (see also Dommergues et al., 1986).

Allometric analyses can provide valuable information about evolutionary modifications of growth trajectories and about patterns of character covariation, whether or not information on age is available. In the absence of age data, adults and immatures of the species at hand can be compared to identify paedomorphosis or peramorphosis. As we have shown, however, allometric patterns do not allow to infer which heterochronic processes produced them. The correlation between size and age is not biologically meaningful, because it is merely a consequence of the fact that both size and age increase monotonically.

Phylogenetic information is essential to establish the direction of heterochronic changes Fink (1982, 1988), and to distinguish general evolutionary trends independently affecting several lineages from innovations appearing locally on a particular branch of a cladogram. As demonstrated by our case study, the analysis of heterochronic processes in such a historic framework can be used to generate hypotheses about the possible ecological background of evolutionary

events (see also Wake and Larson, 1987). By comparing species in a phylogenetic framework, we obtain accounts of evolution that are chronicles of speciation events and character state changes (O'Hara, 1988). Integrating the information about morphological form, development, life history, and phylogeny will help transform these chronicles to an historical narrative which will provide explanations of evolutionary change.

References

Airoldi, J.-P., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia). J. Zool. (London) 216:21–36.

Alberch, P., and J. Alberch. 1981. Heterochronic mechanisms of morphological diversification and evolutionary change in the Neotropical salamander, *Bolitoglossa occidentalis* (Amphibia: Plethodontidae). J. Morphol. 167:249–264.

Alberch, P., S. J. Gould, G. F. Oster, and D. B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5:296–317.

Andersen, N. M. 1982. The Semiaquatic Bugs (Hemiptera, Gerromorpha): Phylogeny, Adaptations, Biogeography, and Classification. Scandinavian Science Press, Klampenborg, Denmark.

Andersen, N. M., and J. R. Spence. 1992. Classification and phylogeny of the Holarctic water strider genus Limnoporus Stål (Hemiptera, Gerridae). Can. J. Zool. 70:753–785.

Andersen, N. M., J. R. Spence, and M. V. H. Wilson. 1993. 50 Million years of structural stasis in water striders (Hemiptera, Gerridae). Am. Entomol. *In press.*

Blackith, R. E., R. G. Davies, and E. A. Moy. 1963. A biometric analysis of development in *Dysdercus fasciatus* Sign (Hemiptera: Pyrrhocoridae). Growth 27:317–334.

Blackstone, N. W. 1987a. Specific growth rates of parts in a hermit crab *Pagurus longicarpus*: A reductionist approach to the study of allometry. J. Zool. (London) 211:531–545.

—. 1987b. Size and time. Syst. Zool. 36:211–215.

Blackstone, N. W., and P. O. Yund. 1989. Morphological variation in a marine hydroid: A comparison of size-based and age-based heterochrony. Paleobiology 15:1–10.

Boughton, D. A., B. B. Collette, and A. R. McCune. 1991. Heterochrony in jaw morphology of needlefishes (Teleostei: Belonidae). Syst. Zool. 40:329–354.

Briggs, J. C. 1987. Biogeography and Plate Tectonics. Elsevier, Amsterdam.

Burnaby, T. P. 1966. Growth-invariant discriminant functions and generalized distances. Biometrics 22:96–110.

Cheverud, J. M., J. J. Rutledge, and W. R. Atchley. 1983. Quantitative genetics of development: Genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905.

Cock, A. G. 1966. Genetical aspects of metrical growth and form in animals. Q. Rev. Biol. 41:131–190.

Cuzin-Roudy, J., and Ph. Laval. 1975. A canonical discriminant analysis of post-embryonic development in *Notonecta maculata* Fabricius (Insecta: Heteroptera). Growth 39:251–280.

Davies, R. G., and V. Brown. 1972. A multivariate analysis of postembryonic growth in two species of *Ectobius* (Dictyoptera: Blattidae). J. Zool. (London) 168:51–79.

Dommergues, J.-L., B. David, and D. Marchand. 1986. Les relations ontogenèse–phylogenèse: Applications paléontologiques. Geobios 19:335–356.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat. Sci. 1:54–77.

Fairbairn, D. J. 1988. Sexual selection for homogamy in the Gerridae: An extension of Ridley's comparative approach. Evolution 42:1212–1222.

—. 1990. Factors influencing sexual size dimorphism in temperate waterstriders. Am. Nat. 136:61–86.

Fink, W. L. 1982. The conceptual relationship between ontogeny and phylogeny. Paleobiology 8:254–264.
—. 1988. Phylogenetic analysis and the detection of ontogenetic patterns, pp. 71–91. *In* M. L. McKinney (ed.), Heterochrony in Evolution: A Multidisciplinary Approach. Plenum Press, New York.
Flury, B. 1988. Common Principal Components and Related Multivariate Models. John Wiley & Sons, New York.
Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biol. Rev. 41:587–640.
—. 1977. Ontogeny and Phylogeny. Harvard University Press, Cambridge, Mass.
Jolicoeur, P. 1963.The multivariate generalization of the allometry equation. Biometrics 19:497–499.
Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.
Klingenberg, C. P., and R. Froese. 1991. A multivariate comparison of allometric growth patterns. Syst. Zool. 40:410–419.
Klingenberg, C. P., and M. Zimmermann. 1992a. Dyar's rule and multivariate allometric growth in nine species of water striders (Heteroptera: Gerridae). J. Zool. (London) 227: 453–464.
—. 1992b. Static, ontogenetic, and evolutionary allometry: A multivariate comparison in nine species of water striders. Am. Nat. 140:601–620.
Laird, A. K., A. D. Barton, and S. A. Tyler. 1968. Growth and time: An interpretation of allometry. Growth 32:347–354.
Lessa, E. P., and J. L. Patton. 1989. Structural constraints, recurrent shapes, and allometry in pocket gophers (genus *Thomomys*). Biol. J. Linn. Soc. 36:349–363.
McKinney, M. L. 1986. Ecological causation of heterochrony: A test and implications for evolutionary theory. Paleobiology 12:282–289.
—. 1988. Classifying heterochrony: Allometry, size and time, pp. 17–34. *In* M. L. McKinney (ed.), Heterochrony in Evolution: A Multidisciplinary Approach. Plenum Press, New York.
Mc Kinney, M. L., and K. J. McNamara. 1991. Heterochrony: The Evolution of Ontogeny. Plenum Press, New York.
McNamara, K. J. 1986. A guide to the nomenclature of heterochrony. J. Paleont. 60:4–13.
—. 1988. The abundance of heterochrony in the fossil record, pp. 287–325. *In* M. L. McKinney (ed.), Heterochrony in Evolution: A Multidisciplinary Approach. Plenum Press, New York.
Mosimann, J. E. 1970. Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions. J. Am. Stat. Ass. 65:930–945.
O'Hara, R. J. 1988. Homage to Clio, or, toward an historical philosophy for evolutionary biology. Syst. Zool. 37:142–155.
Raff, R. A., and G. A. Wray. 1989. Heterochrony: Developmental mechanisms and evolutionary results. J. evol. Biol. 2:409–434.
Reiss, J. O. 1989. The meaning of developmental time: A metric for comparative embryology. Am. Nat. 134:170–189.
Rohlf, F. J., and F. L. Bookstein. 1987. A comment on shearing as a method for "size correction." Syst. Zool. 36:356–367.
Schmidt-Nielsen, K. 1984. Scaling: Why is Animal Size so Important? Cambridge University Press, Cambridge, UK.
Shea, B. T. 1985. Bivariate and multivariate growth allometry: Statistical and biological considerations. J. Zool. (London) 206:367–390.
Sokal, R. R., and F. J. Rohlf. 1981. Biometry. W. H. Freeman Co., New York.
Spence, J. R. 1986. Relative impacts of mortality factors in field populations of the water-strider Gerris buenoi Kirkaldy (Heteroptera: Gerridae). Oecologia (Berlin) 70:68–76.
—. 1989. The habitat templet and life history strategies of pond skaters (Heteroptera: Gerridae): reproductive potential, phenology, and wing dimorphism. Can. J. Zool. 67:2432–2447.

Spence, J. R., and N. M. Andersen. 1994. Biology of water striders: Interactions between systematics and ecology. Ann. Rev. Entomol. *In press*.

Spence, J. R., D. H. Spence, and G. G. E. Scudder. 1980. The effects of temperature on growth and development of water strider species (Heteroptera: Gerridae) of central British Columbia and implications for species packing. Can. J. Zool. 58:1813–1820.

Sperling, F. A. H., and J. R. Spence. 1990. Allozyme survey and relationships of *Limnoporus* Stål Species (Heteroptera: Gerridae). Can. Ent. 122:29–42.

Taylor, F. 1981. Ecology and evolution of physiological time in insects. Am. Nat. 117:1–23.

Wake, D. B., and A. Larson. 1987. Multidimensional analysis of an evolving lineage. Science 238:42–48.

Winterbottom, R. 1990. The *Trimmatom nanus* species complex (Actinopterygii, Gobiidae): Phylogeny and progenetic heterochrony. Syst. Zool. 39:253–265.

Zimmermann, M., R. Hauser, and J. Hüsler. 1982. Schätzung der stadienspezifischen Dichten und Überlebensraten in einer Larvenpopulation von *Gerris lacustris* (Hemiptera, Gerridae). Rev. suisse Zool. 89:941–955.

Zubakov, V. A., and I. I. Borzenkova. 1990. Global Palaeoclimate of the Late Cenozoic. Elsevier, Amsterdam.

5. Ontogeny and Individual Variation: Analysis of Patterned Covariance Matrices with Common Principal Components*

INTRODUCTION

In recent years, there has been renewed interest in the extent to which ontogeny acts as a mechanism influencing patterns of evolutionary change in morphological traits (e.g., Gould, 1977; McKinney and McNamara, 1991; Hall, 1992). Some studies have focused on the outcome of these processes by comparing multivariate patterns of ontogenetic and evolutionary variation (e.g., Shea, 1985; Voss et al., 1990; Klingenberg and Zimmermann, 1992; Voss and Marcus, 1992), whereas others have emphasized microevolutionary processes by studying individual variation in growth and its genetic basis (e.g., Cheverud et al., 1983; Kirkpatrick, 1988; Lynch, 1988; Atchley and Hall, 1991; Cowley and Atchley, 1992; Björklund, 1993).

Numerous studies of the evolution of ontogeny have focused on individual variation in growth curves (also called growth trajectories), i.e., trait measurements as a function of age or developmental stage. Age-specific measurements either can be treated as separate variables, or continuous growth functions can be accommodated by interpolating between the ages at which measurements were made (Kirkpatrick, 1988; Kirkpatrick and Lofsvold, 1989; Kirkpatrick et al., 1990; Björklund, 1993). Usually, the analyses focus on covariances or correlations among measurements made at several ontogenetic stages, but consider only one trait at a time, thereby ignoring correlations among traits (e.g., Cheverud et al., 1983; Leamy and Cheverud, 1984; Lynch, 1988; Kirkpatrick et al., 1990; Björklund, 1993). Cheverud et al. (1983) characterized relationships among variables by first computing the eigenvectors of among-stage covariance matrices for each trait separately, and then comparing them using vector correlations. Björklund (1993) used an analogous procedure within the framework for continuous growth. A more formal approach, however, would simultaneously include all measurements and stages into the analysis. Yet even for moderate numbers of stages and traits, the full statistical model, with no constraints imposed, would contain a very large number of parameters to be estimated, and thus would render the application to real data sets difficult, which is probably why such a study has not been attempted. Nevertheless, a simultaneous analysis of the ontogenies of several traits is feasible if one makes some simplifying assumptions, as they are suggested by the similarity among patterns of ontogenetic variation found in different traits (Cheverud et al., 1983; Björklund, 1993).

Another approach to understanding the connections between ontogeny and evolution focuses on the static variation among individuals at a particular

---

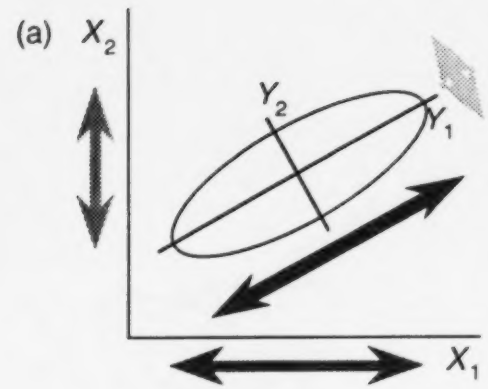
* A version of this chapter has been accepted for publication. Klingenberg, Neuenschwander, and Flury. 1996. Systematic Biology. 45.

stage, which is the raw material upon which natural selection can act. This variation is the product of variation in the developmental processes that generated the structures under study, and can therefore be used to investigate these processes (Cheverud, 1982; Zelditch, 1987; Cowley and Atchley, 1990; Atchley et al., 1992) and their regulation (Tanner, 1963; Atchley, 1984; Riska et al., 1984). Several studies comparing patterns of static variation across ontogenetic stages have found that a single "size" component dominated the variation within each stage (Cuzin-Roudy, 1975; Zelditch, 1988; Klingenberg and Zimmermann, 1992). None of these studies, however, considered the correlations of measurements among stages, either because they were based on cross-sectional data, with a different sample taken independently for each stage, or because the statistical methods were unable to deal with such correlations.

In this article, we introduce a new statistical model (Neuenschwander, 1991; Flury and Neuenschwander, 1995a), which we use to analyze variation in multiple measurements at several ontogenetic stages. The model specifically uses the information contained in longitudinal data, with measurements at all stages for each individual, as it explicitly considers covariation among traits as well as across stages. An extension of the common principal component (CPC) model for independent groups (Airoldi and Flury, 1988; Flury, 1988), the model assumes that CPCs are not only uncorrelated within each group, but also between groups; e.g., the first CPC in one ontogenetic stage is correlated only with the first CPC of other stages. The assumption underlying the CPC model is motivated by the commonalities among traits or stages observed in earlier studies, and our example of growth in the water strider *Limnoporus canaliculatus* demonstrates that it can be realistic. Because CPCs are uncorrelated within and between stages, the model effectively divides a very complex analysis into several simpler ones. Furthermore, it sheds light on the connection between morphometric variability and growth variation, and suggests a coherent framework to study them jointly.

Statistical Models

Principal component analysis (PCA) and its recent extensions are frequently used in morphometric applications, especially in multivariate allometry (Jolicoeur, 1963; Pimentel, 1979; Airoldi and Flury, 1988; Marcus, 1990; Klingenberg, 1996). In most of these applications, the data have a relatively simple structure, and consist either of measurements made on a single group of specimens, or of measurements made on specimens from several separate groups (e.g., species, sexes, ecomorphs or geographical variants). Longitudinal growth data, however, have a more complex structure, because the same individuals are measured for each growth stage, and the stages therefore cannot be treated as independent groups. This interdependence requires substantial adjustments of the statistical models used to analyze such data. Because multivariate studies of multiple groups and longitudinal studies are based on rather complex data, the need to summarize these data using simplified models is especially urgent. In this section, we briefly review one-group PCA and models of common principal components (CPCs) for independent groups before we introduce patterned covariance matrices for longitudinal data and a model of CPCs for dependent random vectors. We emphasize the use of principal component (PC) and CPC models as tools for data reduction.



(b)

	X_1	X_2
X_1	7.1	3.4
X_2	3.4	2.9

(c)

	Y_1	Y_2
Y_1	9	0
Y_2	0	1

Fig. 5-1. Principal component analysis of a single group. (a) Bivariate plot with a contour ellipse representing the distribution of data points. The original variables X_1 and X_2 are transformed into the principal components Y_1 and Y_2, which account for maximal and minimal variation, respectively (arrows). This transformation corresponds to a rotation of the coordinate system, in which the new coordinates are aligned with the major and minor axes of the contour ellipse. (b) Covariance matrix of the original variables. Units are arbitrary. (c) Covariance matrix of the principal components. Notice the diagonal structure of this matrix, i.e., the off-diagonal elements of the matrix are zero values. Units are arbitrary.

One-Group Principal Components and Common Principal Components for Independent Groups

PCA is a tool to analyze variation within a single group of specimens. In the space spanned by the variables (e.g., in two dimensions the plane of a scatter plot), PCA can be used to assess the amount and direction of this variation. It transforms the original variables into PCs, a set of new variables that successively account for the largest possible part of total variation while being uncorrelated among each other (Fig. 5-1a). This transformation fundamentally changes the covariance matrix (Fig. 5-1b), and renders it into diagonal form (Fig. 5-1c): because PCs are uncorrelated, all off-diagonal elements of the covariance matrix (covariances between pairs of PCs) are zero, whereas the diagonal contains the variances of the PCs, or eigenvalues (see Appendix, PCA).

In many applications, the first few PCs account for the largest portion of total variance. In morphometrics, it is not uncommon that the first one or two PCs take up 95% or more of the variation in a much larger number of variables. The first few PCs therefore summarize most of the variation in fewer dimensions, perhaps only one. Models of this kind, including only one allometric "size" axis while regarding the remaining variation as random scatter around it, have been used traditionally in morphometrics (e.g., Hopkins, 1966; Bookstein et al., 1985; Klingenberg, 1996). This extreme data reduction using simplified models of within-group variation is helpful for comparisons between two or more groups of specimens.

PCA has been generalized for situations involving several groups. The CPC model assumes that the groups all share the same (common) PCs, but it allows the groups to differ in the amounts of variation associated with each one (Airoldi and Flury, 1988; Flury, 1988). The scatter ellipsoids for all groups therefore have parallel principal axes, but the lengths of corresponding axes may vary. Under the CPC model, a single transformation simultaneously converts the covariance matrices of all groups to diagonal form (Appendix, CPCs for Independent Groups). Applications of the CPC model to biological data sets include Airoldi and Flury (1988), Klingenberg and Zimmermann (1992), and Klingenberg and Spence (1993).

Longitudinal Data and Patterned Covariance Matrices

The CPC model introduced above was designed for independent groups, e.g., samples drawn from different sexes, ecomorphs, or species. Numerous growth studies use separate samples of specimens in different ontogenetic stages; such cross-sectional data can be analyzed using the CPC model for independent groups (e.g., Klingenberg and Zimmermann, 1992). Longitudinal data, however, consist of measurements made on the same specimens in several growth stages (Fig. 5-2a), and these stages therefore are not independent groups. Following individuals through growth has obvious benefits, because it allows to address questions about individual variation in growth processes and their regulation, e.g., whether there is compensatory growth (Tanner, 1963; Monteiro and Falconer, 1966; Riska et al., 1984; Lynch, 1988; Kirkpatrick et al., 1990; Cowley and Atchley, 1992). As a consequence of this additional information, however, longitudinal data have a more complex structure than cross-sectional data.

For a typical longitudinal study, in each of k different growth stages p measurements are taken on the same n specimens (Fig. 5-2a). The data are most conveniently arranged in a $n \times kp$ matrix, i.e., the measurements for the different stages are treated as separate variables. The resulting covariance matrix has a distinctive pattern: it consists of an array of $k \times k$ submatrices, each of dimension $p \times p$ (Fig. 5-2b). The blocks along the main diagonal of the matrix (shaded gray in Fig. 5-2b) are the within-stage covariance matrices, as they also are used in cross-sectional analyses. They characterize static variation between individuals within each stage. The off-diagonal blocks (diagonal hatching and arrows in Fig. 5-2b) contain covariances between measurements taken in different growth stages.

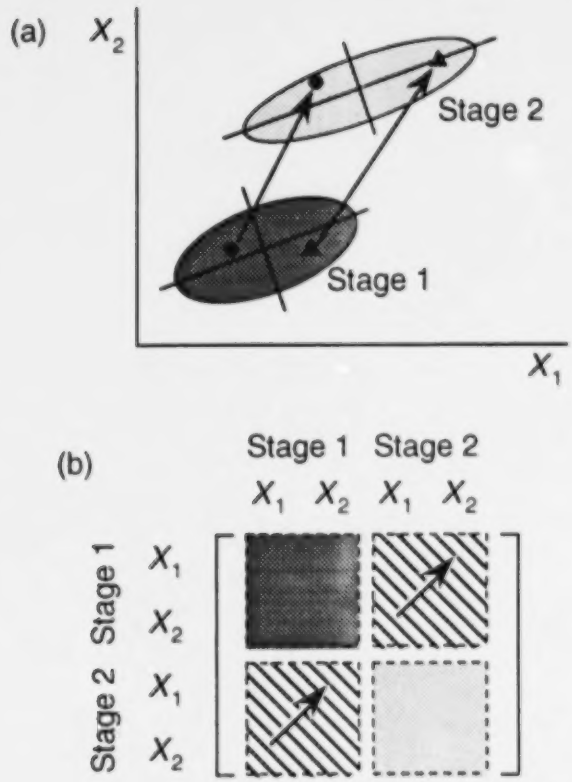


Fig. 5-2. The structure of longitudinal data. (a) Bivariate scatter plot with contour ellipses for two growth stages (i.e., $p = 2$, $k = 2$). The measurements for two specimens (dots and triangles) are plotted in each stage and connected by arrows. (b) Patterned covariance matrix describing the data. The morphometric measurements X_1 and X_2 of each stage are entered separately as variables for the analysis. The blocks on the main diagonal (shaded) are the within-stage covariance matrices of stages 1 and 2. The off-diagonal blocks (hatched, with arrows) contain covariances of measurements between stages, which can be used to study regulatory phenomena such as compensatory growth.

The number of variables in a longitudinal analysis can be very large, even with moderate numbers of measurements and stages. The example we use to demonstrate this approach (see below) contains four measurements and six growth stages, and is thus smaller than the data sets used in many similar studies. Nevertheless, this means there are 24 variables in the analysis, and without further constraints, the number of parameters required for a full statistical description of the covariance structure is 300 ($=24 \times [24 + 1] / 2$; see Appendix). This complexity of longitudinal data calls for techniques of data reduction; therefore, we introduce a simplified model, for which fewer parameters need to be estimated, but which still can represent the data with reasonable accuracy.

Common Principal Components for Dependent Random Vectors

The CPC model presented above assumes that a single transformation simultaneously converts the covariance matrices of k groups to diagonal form. Because the groups are assumed to be independent of one another, these covariance matrices characterize the variation within groups sufficiently. If the

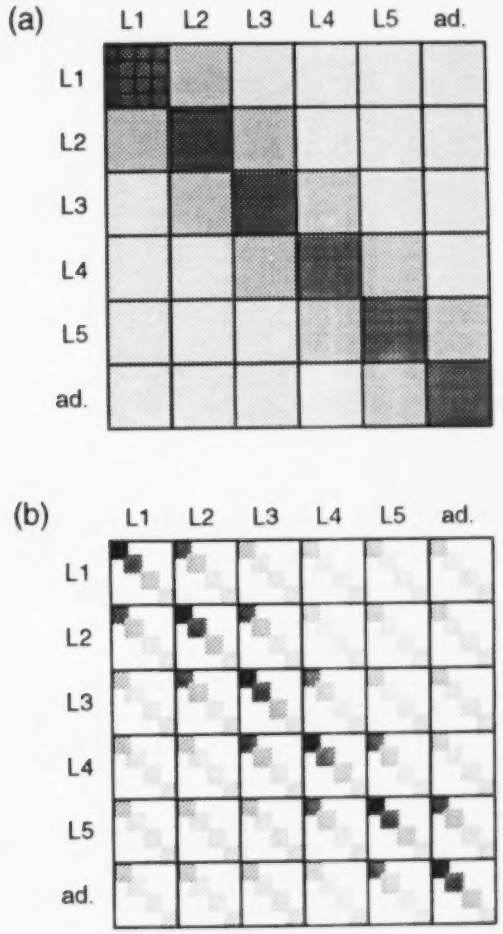


Fig. 5-3. The model of common principal components for dependent random vectors. The structure of longitudinal data is shown as in the water strider example, with six discrete growth stages (five larval instars, denoted L1–L5, and the adult stage) and four morphometric variables. Shading represents the approximate magnitude of matrix elements (blank for zero). (a) Diagram of the covariance matrix. The blocks along the diagonal are the covariance matrices within stages, whereas the off-diagonal blocks contain covariances between stages. Each block is a 4×4 matrix, as indicated by the grid in the covariance matrix for the L1. (b) Covariance matrix of the CPCs. All the submatrices have diagonal form, because only pairs of corresponding PCs are correlated between instars.

groups are interdependent, however, there are k sets of measurements (groups, growth stages), each with the same p variables, for every observation (examples in Flury and Neuenschwander, 1995a). The result is a patterned $kp \times kp$ covariance matrix composed of k^2 submatrices (each of format $p \times p$): the k within-group covariance matrices are arranged as blocks along the diagonal, while the off-diagonal blocks are matrices of covariances between groups. Figure 5-3a shows such a patterned covariance matrix for ontogenetic data, where the groups correspond to discrete growth stages.

Like the CPC model for independent groups, the model of CPCs for dependent random vectors assumes that all the groups share the same PCs. Therefore, the transformation to CPCs converts all the within-group covariance matrices to diagonal form (the blocks along the diagonal in Fig. 5-3b). In addition, however, the same transformation must also render all the remaining submatrices diagonal, which contain the covariances of measurements across groups (Appendix, CPCs for Dependent Random Vectors; Neuenschwander, 1991; Flury and Neuenschwander, 1995a). This means that only corresponding CPCs are correlated among groups; for example, only the pairs of first or of second CPCs are correlated among groups, but not the first CPC in one group with the second CPC in another group.

This CPC model results in a substantial reduction of the number of parameters that have to be estimated. In the example with $k = 6$ and $p = 4$, there are only 90 parameters instead of 300 in the unconstrained model (see Appendix, Number of Parameters). The advantage of the model becomes more apparent if the covariance matrix shown in Figure 5-3b is rearranged so that rows and

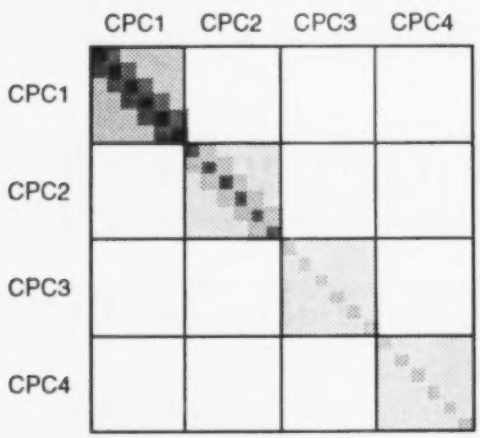


Fig. 5-4. Rearrangement of the patterned covariance matrix after CPC transformation (cf. Fig. 5-3b). The transformation considerably simplified the structure of longitudinal growth data. Because the CPCs are uncorrelated both within and among stages, they can be examined separately. The multivariate problem is thus reduced to a study of the matrix of covariances between stages for each CPC. If some CPCs only account for a small amount of variation, they may be omitted from the analysis.

columns are ordered by CPCs rather than by groups (Fig. 5-4). This rearranged matrix consists of p^2 blocks of format $k \times k$. All the elements of the off-diagonal blocks are zero values. The CPCs can now be studied separately because they are uncorrelated among each other.

For longitudinal growth studies with multiple measurements, this model therefore reduces a complex multivariate problem into a number of simpler analyses, each considering one CPC. Thus one has to examine p matrices of covariances of CPC scores among developmental stages, using the methods developed for analyses of a single measurement. Moreover, CPCs accounting for only minor amounts of variation may be ignored in the interpretation of results, as in one-group PCA, to simplify the analysis even further.

Example: Growth in Water Striders

Data

Water striders (Heteroptera: Gerridae) are especially suitable for studying the ontogeny of individual variation, because their growth occurs in six discrete stages: five larval instars, denoted L1–L5, precede the adult stage. Because there is no variation in the number of larval instars, they are comparable developmental stages. Due to the rigid cuticle, the growth of numerous structures can be followed easily for the entire postembryonic development, as in other hemimetabolous insects. Moreover, water striders can be reared individually in the laboratory, and it is easy to obtain a complete record of each specimen's growth by collecting the exuviae at every molt and the adult.

For this study, we used longitudinal growth data from the water strider *Limnoporus canaliculatus* (Say), reared under controlled laboratory conditions (20°C, photoperiod 16L:8D). Water striders collected in the wild (Morris County, New Jersey; May 1, 1992) were set up as a mass culture, from which eggs were taken for individual rearings. Within about 12 hr of hatching, first instar larvae were put separately into plastic containers (diameter 11.5 cm, height 8 cm), each with about 1 cm of water and a small Styrofoam strip floating on the surface. Each bug was fed a frozen flesh fly [*Neobelliera bullata* (Parker)] daily and checked for molts at intervals of about 12 hr, and all exuviae were collected. After adult emergence, bugs were killed by freezing. Exuviae and adults were stored in 70% ethanol for several months before measuring.

The variables used in this study are the lengths of the femora and tibiae of the middle and hind legs, measured on both body sides. Shrinking and other artifacts due to preservation are negligible, because the cuticle of the legs consists of rigidly sclerotized tubes even in the otherwise delicate first instar exuviae. All measurements were taken with a video system attached to a dissecting microscope.

If the measurements could be made on both body sides, we used arithmetic means of left and right sides; otherwise we used measurements from one side. Data were checked for outliers for each variable separately in every instar. A few individuals had to be excluded from the data set because of deformities related to abnormal molting. In this study, we use the data for 89 females for

which complete data were available in each instar (the data set is available from C.P.K. on request). All measurements were transformed to natural logarithms before the analysis.

Statistical Analysis and Results

The covariance matrix of stage-specific measurements is the basis for the following analysis. The most conspicuous feature of this matrix is a general increase in variances from early to late instars, but especially in the L5 and adult instars (Fig. 5-5). The tibia of the hind legs is the more variable trait in each instar. This pattern is not only repeated in every instar, but to a certain extent it also applies to the covariances among instars; the division of the covariance matrix into blocks is therefore visible mostly from the "peaks" representing the

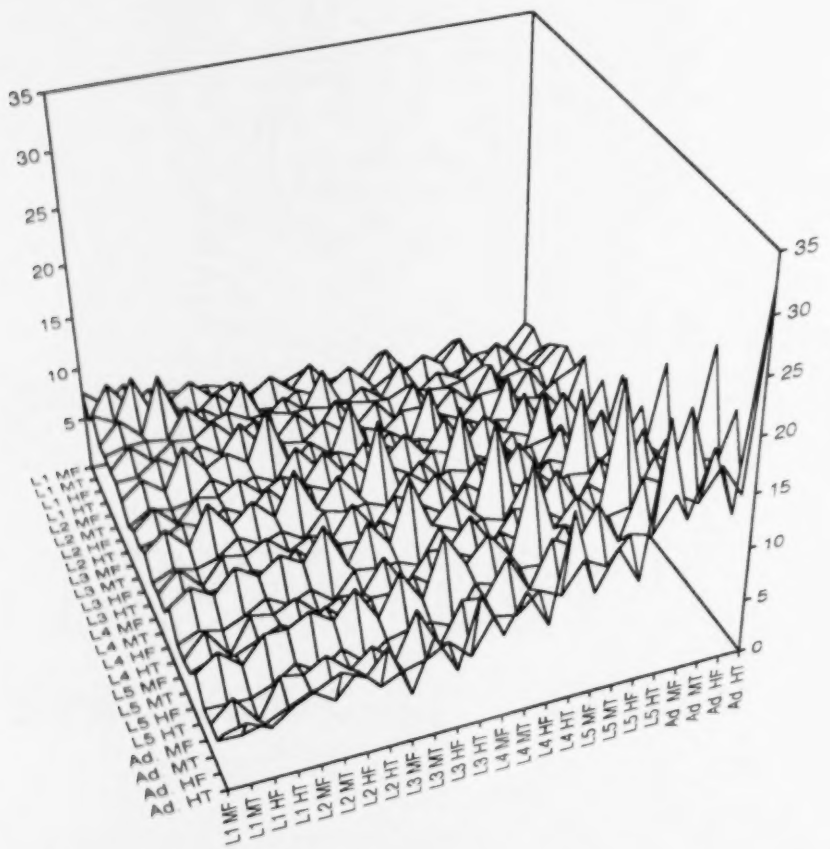


Fig. 5-5. Covariance matrix for stage-specific measurements in females of the water strider *Limnoporus canaliculatus*. The lengths of middle and hind femora and tibiae (MF, MT, HF, and HT) are the four measurements taken for all five larval instars (L1–L5) and adults (Ad.). Values on the vertical axis are variances and covariances for natural log-transformed measurements, multiplied by 10^4.

Table 5-1. Patterns of individual variation in water striders, analyzed by separate PCAs in each instar. Values presented are the PC coefficients for each variable and the percentages of total variance (% variance) for which each PC accounts.

Instar	Variable[a]	PC1	PC2	PC3	PC4
L1	MF	0.455	0.342	–0.144	0.809
	MT	0.500	–0.236	0.832	–0.034
	HF	0.516	0.612	–0.159	–0.578
	HT	0.526	–0.672	–0.511	–0.102
	% variance	76.9	14.8	5.1	3.2
L2	MF	0.504	0.434	0.217	–0.714
	MT	0.486	–0.128	0.716	0.484
	HF	0.482	0.489	–0.556	0.468
	HT	0.527	–0.745	–0.361	–0.191
	% variance	80.7	11.7	4.8	2.8
L3	MF	0.445	0.532	0.128	–0.709
	MT	0.496	–0.167	0.787	0.328
	HF	0.457	0.510	–0.426	0.591
	HT	0.589	–0.656	–0.428	–0.199
	% variance	77.4	14.9	4.1	3.7
L4	MF	0.444	0.497	0.059	–0.743
	MT	0.481	–0.057	0.816	0.315
	HF	0.448	0.514	–0.450	0.576
	HT	0.609	–0.697	–0.357	–0.131
	% variance	80.2	13.4	4.5	1.9
L5	MF	0.436	0.465	–0.095	0.765
	MT	0.474	0.058	0.856	–0.199
	HF	0.446	0.499	–0.423	–0.610
	HT	0.622	–0.729	–0.282	0.054
	% variance	80.3	13.0	4.8	1.8
Ad.	MF	0.430	0.462	–0.138	0.763
	MT	0.513	0.051	0.840	–0.169
	HF	0.422	0.513	–0.414	–0.623
	HT	0.611	–0.722	–0.322	0.034
	% variance	79.4	14.3	4.8	1.5

[a] MF = middle femur; MT = middle tibia; HF = hind femur; HT = hind tibia.

hind tibia in different instars (Fig. 5-5). All covariances within and between instars are positive, and covariances tend to be higher between consecutive instars than between stages that are farther apart.

The patterns of variation within instars can be characterized separately with PCA. Despite the general increase in the amount of variation from instar to instar (Fig. 5-5), the proportion for which each PC accounts remains fairly con-

stant. Within each instar, the PC1 accounts for the largest proportion of the total variance, and the PC1s have coefficients that are all positive (Table 5-1). Therefore, they can be interpreted as "size" vectors reflecting static allometry. The allometric patterns are similar in all instars, as indicated by the high vector correlations among the PC1s, which all exceed 0.99, and the corresponding angles, which range from 0.97° to 7.5°. The PC2, which takes up a moderate amount of variation in all instars, is a contrast of the hind tibia against the middle and hind femora; the middle tibia has coefficients of smaller magnitude, which even vary in their sign. The PC3 and PC4 only account for small proportions of the total variation.

CPCs for dependent random vectors.—Whereas one-group PCA always can transform the covariance matrix of a sample to exactly diagonal form, simultaneous analyses of multiple groups generally pose more difficult problems. Even if the CPC model holds, sampling variation will generally make it impossible for any single transformation to render all blocks of the covariance matrix perfectly diagonal. Estimating the CPC coefficients is therefore an optimization process, searching for a transformation that minimizes a measure of deviation from simultaneous diagonality in all blocks of the covariance matrix (Appendix, Estimation of CPCs). For this purpose, we used a version of the orthogonal *FG*$^{+}$ algorithm (Neuenschwander, 1991; Flury and Neuenschwander, 1995b) written in the SAS/IML language (a text file with this routine is available through the Internet: file://life.bio.sunysb.edu/morphmet/dcpc.exe.ibmpc; a version written in GAUSS is available from B.E.N. or B.D.F.). Standard errors for the estimates of the CPC coefficients were computed using the jackknife method (leave-one-out procedure; e.g., Efron and Tibshirani, 1993: chapter 11). The ordering of CPCs is somewhat arbitrary, because the amounts of variation they take up in each instar can vary. Here, we ranked them according to the average proportion of total variance or between-instar covariance for which they accounted in each block of the covariance matrix.

The CPC coefficients (Table 5-2) closely correspond to the patterns seen in the PCAs for single instars. The CPC1 is an overall "size" axis, weighting the hind tibia somewhat more and the middle tibia slightly less than the the two

Table 5-2. Joint pattern of variation in all six instars of the water strider data. Values presented are CPC coefficients (and their jackknife standard errors).

Variable[a]	CPC1	CPC2	CPC3	CPC4
MF	0.471	0.436	–0.059	0.765
	(0.020)	(0.042)	(0.092)	(0.023)
MT	0.414	0.073	0.878	–0.229
	(0.024)	(0.051)	(0.035)	(0.107)
HF	0.491	0.466	–0.427	–0.600
	(0.028)	(0.043)	(0.087)	(0.059)
HT	0.605	–0.767	–0.209	0.048
	(0.033)	(0.031)	(0.051)	(0.051)

[a] MF = middle femur; MT = middle tibia; HF = hind femur; HT = hind tibia.

femur lengths, whereas the other CPCs are contrasts between measurements that are fairly similar to the corresponding within-instar PCs. We applied the CPC transformation to all six instars, and thereby changed the covariance matrix of the original variables (Fig. 5-5) into that of the CPCs (Fig. 5-6; cf. Figs. 5-3a, 5-3b). The covariance matrix of transformed variables is an array of "spikes" and thus shows that the CPC1 clearly dominates the variation in all instars, and also accounts for almost all the covariance among instars (Fig. 5-6). The percentages of within-instar variance taken up by the CPC1 are only slightly lower than those for which the PC1s account (cf. Tables 5-1, 5-3). Therefore, the CPC1 is a fairly good summary of static allometry in all instars jointly.

Rearranging the covariance matrix by CPCs makes the dominance of the CPC1 even more visible (Fig. 5-7). The variance in this "size" component is

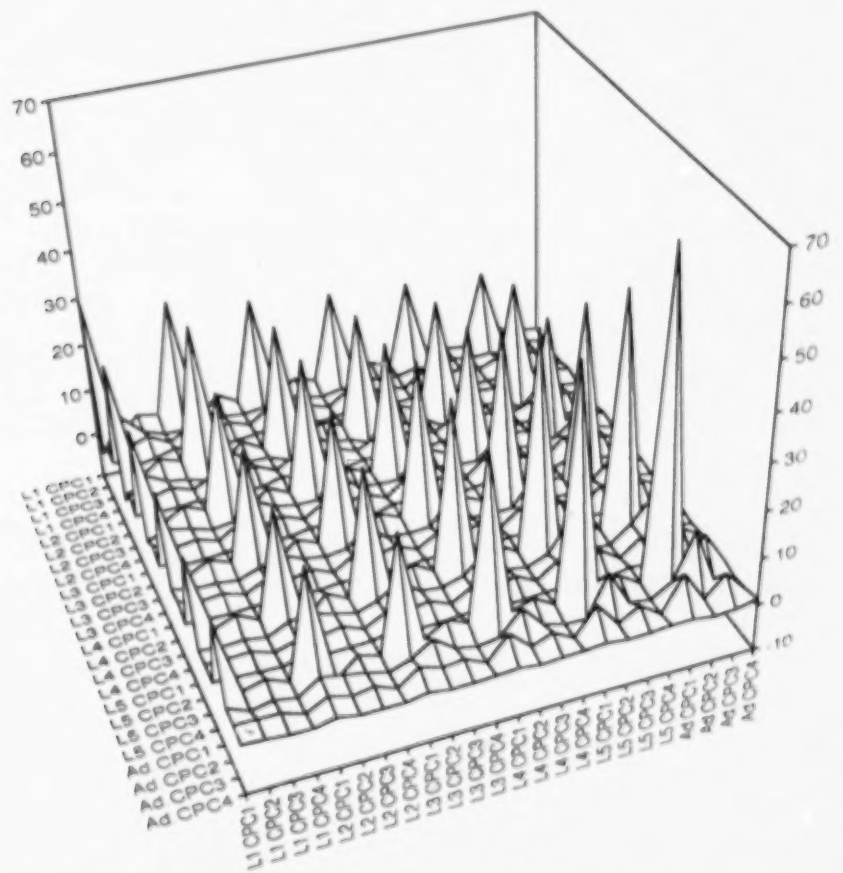


Fig. 5-6. Covariance matrix after transformation to CPCs within each instar. "Spikes" are produced by the variances and between-instar covariances of the first CPC. Values on the vertical axis are variances and covariances of CPCs for natural log-transformed measurements, multiplied by 10^4.

TABLE 5-3. Percentages of total within-instar variance taken up by each CPC of the water strider data.

Instar	CPC1	CPC2	CPC3	CPC4
L1	75.9	13.8	6.7	3.6
L2	79.7	11.9	5.3	3.1
L3	76.8	14.0	5.4	3.9
L4	79.7	13.2	5.1	1.9
L5	79.8	13.2	5.2	1.8
Adult	78.2	14.4	5.8	1.5

fairly constant in the L1–L3, but later increases markedly from instar to instar. In the CPC1 and CPC2 the covariances among instars are highest between successive instars, and decline substantially as the number of intervening instars increases. The variance of the CPC2 remains fairly constant in the younger instars and gradually increases in the L5 and final instars. As this increase parallels that in the CPC1, the CPCs account for similar proportions of total variance in each instar (Table 5-3).

Most of the covariances between different CPCs (off-diagonal blocks in Fig. 5-7) are low, suggesting that the CPCs are almost uncorrelated. There are, however, weak to moderate positive correlations between the CPC1 and CPC3, ranging from 0.12 to 0.41, and between the CPC2 and CPC3 (0.07–0.32). Negative correlations are especially frequent between the CPC4 scores in the L1–L4 and the CPC2 and CPC3. The other correlations between CPCs do not display any apparent pattern and most are substantially weaker (total range –0.27 to 0.23). This indicates that the CPC model fits these data fairly well.

Test of the CPC model.—To evaluate whether these correlations seriously violate the assumptions of the CPC model, we used permutation tests, also known as randomization tests (Edgington, 1986, 1987; Manly, 1991; Efron and Tibshirani, 1993: chapter 15; Westfall and Young, 1993; Good, 1994). This class of tests uses repeated permutations of the original data to simulate the distribution of a test statistic under the null hypothesis stating that two or more samples are drawn from the same population or that several variables are uncorrelated. Our test is based on the fact that, under the CPC model, different CPCs are uncorrelated within and between instars (Figs. 5-3b, 5-4). It follows the procedure for testing bivariate correlations against the null hypothesis of independence by reshuffling the values of one variable repeatedly (Pitman, 1937; Edgington, 1987:198–201). We first computed the CPC scores for each individual, and then randomly reshuffled the observations separately for the CPC2–CPC4, each time keeping all instars together. This left the associations among instars unchanged for each CPC, because the permutation procedure affected only the correlations between different CPCs. This step was repeated 1,000 times. We calculated the CPCs for each of the randomized data sets, and computed three different test statistics: (1) the e statistic of deviation from simultaneous diagonality in all blocks of the covariance matrix (Appendix, Estimation of CPCs) as an overall test, (2) the maximum absolute covariance, and

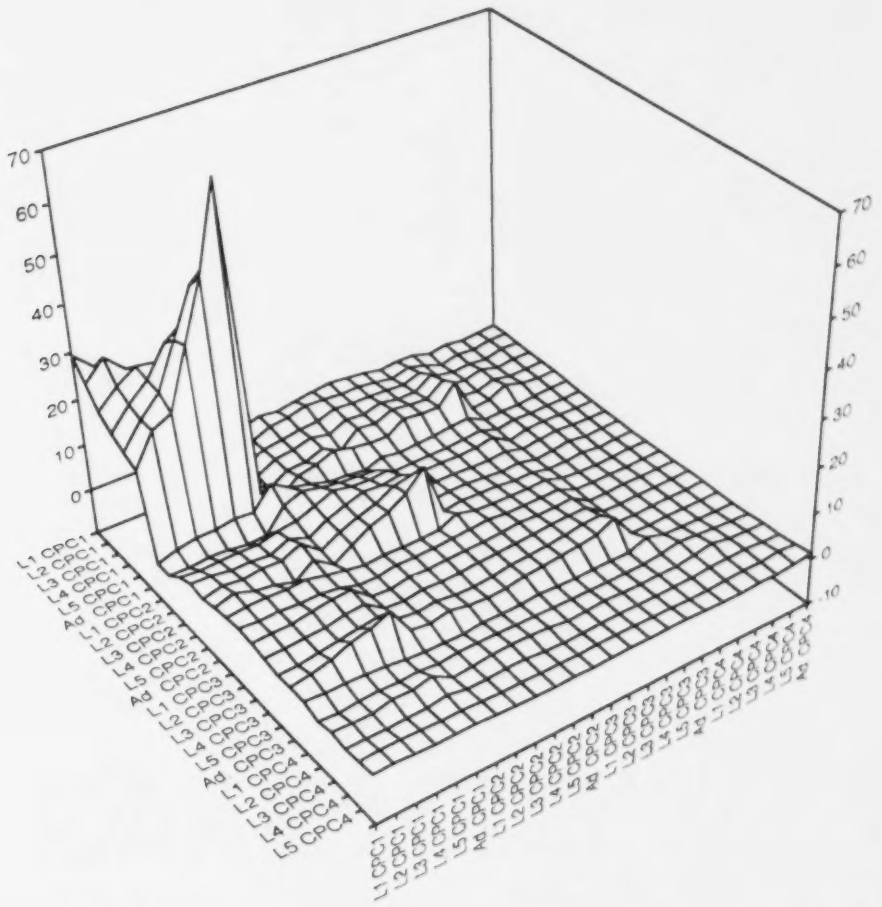


Fig. 5-7. Covariance matrix arranged by CPCs. The covariation of the CPCs among instars is now more apparent here than in Figure 5- 6. Values on the vertical axis are variances and covariances of CPCs for natural log-transformed measurements, multiplied by 10^4.

(3) correlation between different CPCs (i.e., excluding the diagonal entries in each block of the covariance matrix).

The *e* statistic did not reach the observed value in any of the 1,000 simulation runs of the CPC model, and therefore provides strong evidence that the CPC model does not fit the data well overall. The maximal covariance and correlation between different CPCs of the original data, both between the CPC1 and CPC3 in adults (see Fig. 5-7), were attained in only six or eight randomization runs, and thus supported the result obtained with the *e* statistic. Because the CPC3 accounts for a small fraction of the total variation, despite its relatively high correlations with other CPCs, we repeated the randomization test with two new statistics: the maximal absolute covariance and correlation not involving the CPC3. These statistics matched or exceeded the observed values

in 40.1% and 34.2% of the randomization runs, respectively. From this we conclude that the statistically significant deviations from the CPC model are related to the CPC3, but neither covariances nor correlations among the other CPCs are distinguishable from random effects.

DISCUSSION

In this article, we have introduced a statistical model designed to analyze the variation of multiple variables in several interdependent groups. Our example is an application of this method to the familiar problem of longitudinal data with a number of measurements taken in several ontogenetic stages (e.g., Cuzin-Roudy, 1975; Cheverud et al., 1983; Björklund, 1993). Because statistical methods to deal with such a complex data structure have not been available, previous authors either had to treat different stages independently and thus ignore the longitudinal nature of the data, or they were forced to perform the analyses separately for each measurement, thereby neglecting the correlations among traits. A model that specifically addresses the complexity of the data structure offers several advantages in this situation. First of all, it allows to include all traits in a single analysis rather than to compare the results from separate analyses in an informal manner. Moreover, the simplifying assumptions made by the CPC model, when met, can provide further insight into the underlying patterns of variation and they can lead to substantial data reduction if most variation can be approximated by just a few CPCs. Below, we will illustrate these points with the results from our example.

First, however, we have to assess the fit of the model to the data set. The model assumes that the CPCs are independent of each other both within and across developmental stages, and that nonzero covariances and correlations between different CPCs are due to sampling variation. Covariances between CPCs are generally low (Fig. 5-7), although there are relatively high correlations involving the CPC3. Unfortunately, tests based on large-sample theory (Neuenschwander, 1991; Flury and Neuenschwander, 1995a) are not reliable in this case, because the sample size, 89, is fairly small compared to the 24 variables in the model (i.e., four measurements in each of six stages). For this reason, we used a permutation test (Pitman, 1937; Edgington, 1986, 1987; Manly, 1991; Westfall and Young, 1993; Good, 1994). Overall, there are significant deviations from the CPC model, but closer examination showed that they all concern the CPC3, which accounts for only a minor proportion of the total variation. The other CPCs are nearly uncorrelated within and across instars, and with some caution, therefore, the CPC model can be applied for these.

Using the CPC model, although it may not fit perfectly to the data, has a major benefit because it dramatically simplifies the problem by reducing the number of parameters to be estimated. In our example, the CPC model uses less than a third of the parameters it takes for a full statistical description with the unconstrained model. The advantage becomes even more tangible if one realizes that only these simplifying assumptions made longitudinal growth data with multiple measurements statistically tractable. Unlike the original 24×24 covariance matrix (Fig. 5-5), the transformed and rearranged matrix (Fig. 5-7) shows some simple patterns, for which biological interpretations can be sought. This benefit far outweighs the relatively minor misfit of the model.

The CPC model can also be an effective tool for further data reduction. In our example, the CPC3 and CPC4 account for small amounts of variation in all instars, and probably can be ignored for most purposes (see Table 5-3). The CPC2 has a moderate degree of variability in all instars (Fig. 5-7), and to give a fairly complete description of morphometric variability throughout ontogeny, it should be considered along with the CPC1.

The CPC1, which is an "overall size" component, takes up the largest proportion of static variability within each stage and also accounts for most of the covariance between instars. The variances of the CPC1 are fairly constant in the L1–L3 instars, but then increase from instar to instar. This suggests that variability in "size" added at each molt is first compensated by some regulatory mechanism, which is switched off in the L4, leading to divergent growth in the later instars (Riska et al., 1984). Such variability in growth regulation between ontogenetic stages has also been shown in other arthropods (e.g., Hartnoll and Dalley, 1981; Tanaka, 1981; West and Costlow, 1987). That growth is not strongly "targeted" is further supported by the covariances of the CPC1 between instars, which are all positive and relatively high (correlations are 0.37–0.93), indicating that individuals tend to be either relatively small or relatively large in all instars (a detailed analysis will be presented elsewhere; Klingenberg, unpubl.). The CPC2 and CPC3 take up moderate or small amounts of variance throughout the entire life cycle.

The basis of the CPC model is an assumption about the covariation of morphometric traits: all ontogenetic stages share the same structure of static variation, which also forms the pattern of covariation among instars. As a consequence of this parallelism, the CPCs are uncorrelated not only within each stage, but also among stages. Therefore, variation of the CPCs during growth can be studied separately, using the methods developed for single traits (e.g., Cheverud et al., 1983; Lynch, 1988; Kirkpatrick and Lofsvold, 1989). This procedure, although superficially similar to the approach with a separate analysis for each measurement, does not ignore the correlations among traits, because the CPCs explicitly account for them. Therefore, the CPC model divided a very complex analysis into a few simpler ones. A quick comparison of the covariance matrices shown in Figures 5-5 and 5-7, which contain the same information as both CPC transformation and rearranging are reversible, demonstrates the effectiveness of this approach.

Morphometric variation within stages has long been the focus of numerous studies under the headings of static allometry (e.g., Cuzin-Roudy, 1975; Gibson et al., 1984; Klingenberg and Zimmermann, 1992) and morphological integration (e.g., Cheverud, 1982, 1995; Leamy and Atchley, 1984; Zelditch, 1987, 1988; Wagner, 1990). Although these studies differ widely in their goals and methods used, they all examine the strength of associations among traits. In the CPC framework, the dominance of the CPC1 reflects these associations. Because this pattern applies to covariances between as well as within instars (Fig. 5-6), the CPC approach extends the study of integration from isolated examinations of within-stage variability to a unified analysis of growth. Such an analysis can also investigate the variability and possible evolutionary constraints of growth curves by studying the covariances of each CPC between stages (Kirkpatrick and Lofsvold, 1992; Björklund, 1993).

Covariances among traits in different ontogenetic stages indicate variation and possible constraints for the evolution of ontogenetic trajectories. This

unique information is only available from longitudinal studies. Nevertheless, because such studies are very labor-intensive and only feasible for organisms that can be reared in the laboratory (but see Björklund, 1993), few such studies exist. Cross-sectional studies comparing the ontogenies of different species (e.g., Klingenberg and Spence, 1993) or patterns of allometry (Klingenberg and Zimmermann, 1992) are alternative approaches to these problems. A combination of all these methods is most promising to provide an integrated understanding of individual variation, growth, and evolution.

The CPC model demonstrates the intimate link between the comparison of static variation in several stages and the study of variability in growth curves, which have been the two most important approaches to the study of the connection between ontogenetic processes and evolution. We believe that this model, applied to phenotypic or genetic covariance matrices, will be a useful tool to further explore this connection.

REFERENCES

Airoldi, J.-P., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia). J. Zool. Lond. 216:21–36.

Atchley, W. R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. Am. Nat. 123:519–540.

Atchley, W. R., D. E. Cowley, C. Vogl, and T. McLellan. 1992. Evolutionary divergence, shape change, and genetic correlation structure in the rodent mandible. Syst. Biol. 41:196–221.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. Biol. Rev. 66:101–157.

Björklund, M. 1993. Phenotypic variation of growth trajectories in finches. Evolution 47:1506–1514.

Bookstein, F. L., B. Chernoff, R. L. Elder, J. M. Humphries, Jr., G. R. Smith, and R. E. Strauss. 1985. Morphometrics in evolutionary biology: The geometry of size and shape change, with examples from fishes. Spec. Publ. Acad. Nat. Sci. Phila. 15:i–xvii, 1–277.

Cheverud, J. M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

Cheverud, J. M. 1995. Morphological integration in the saddle-backed tamarin (*Saguinus fuscicollis*) cranium. Am. Nat. 145:63–89.

Cheverud, J. M., J. J. Rutledge, and W. R. Atchley. 1983. Quantitative genetics of development: Genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905.

Cowley, D. E., and W. R. Atchley. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. Am. Nat. 135:242–268.

Cowley, D. E., and W. R. Atchley. 1992. Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. Evolution 46:495–518.

Cuzin-Roudy, J. 1975. Étude de la variabilité et de l'allométrie de taille chez *Notonecta maculata* Fabricius (Insectes, Hétéroptères), par les méthodes classiques et par la méthode des composantes principales. Arch. Zool. Exp. Gén. 116:173–227.

Edgington, E. S. 1986. Randomization tests. Pages 530–538 *in* Encyclopedia of statistical sciences, Volume 7 (S. Kotz and N. L. Johnson, eds.). Wiley, New York.

Edgington, E. S. 1987. Randomization tests, 2nd ed. Dekker, New York.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Flury, B. 1988. Common principal components and related multivariate models. Wiley, New York.

Flury, B. D., and B. E. Neuenschwander. 1995a. Principal component models for patterned covariance matrices, with applications to canonical correlation analysis of several sets of variables. Pages 90–112 *in* Recent advances in descriptive multivariate analysis (W. J. Krzanowski, ed.). Oxford Univ. Press, Oxford, England.

Flury, B. D., and B. E. Neuenschwander. 1995b. Simultaneous diagonalization algorithms with applications in multivariate statistics. Pages 179–206 *in* Approximation and computation: A festschrift in honor of Walter Gautschi (R. Zahar, ed.). Birkhäuser, Boston.

Gibson, A. R., A. J. Baker, and A. Moeed. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): An application of the jackknife to principal component analysis. Syst. Zool. 33:408–421.

Good, P. 1994. Permutation tests: A practical guide to resampling methods for testing hypotheses. Springer-Verlag, New York.

Gould, S. J. 1977. Ontogeny and phylogeny. Harvard Univ. Press, Cambridge, Massachusetts.

Hall, B. K. 1992. Evolutionary developmental biology. Chapman and Hall, London.

Hartnoll, R. G., and R. Dalley. 1981. The control of size variation within instars of a crustacean. J. Exp. Mar. Biol. Ecol. 53:235–239.

Hopkins, J. W. 1966. Some considerations in multivariate allometry. Biometrics 22:747–760.

Jackson, J. E. 1990. A user's guide to principal components. Wiley, New York.

Jobson, J. D. 1992. Applied multivariate data analysis. Volume II. Categorical and multivariate methods. Springer-Verlag, New York.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. Biometrics 19:497–499.

Jolliffe, I. T. 1986. Principal component analysis. Springer-Verlag, New York.

Kirkpatrick, M. 1988. The evolution of size in size-structured populations. Pages 13–28 *in* Size-structured populations (B. Ebenmann and L. Persson, eds.). Springer-Verlag, Berlin.

Kirkpatrick, M., and D. Lofsvold. 1989. The evolution of growth trajectories and other complex quantitative characters. Genome 31:778–783.

Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.

Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124:979–993.

Klingenberg, C. P. 1996. Multivariate allometry. Pages (00)–(00) *in* Advances in morphometrics (L. F. Marcus, M. Corti, A. Loy, D. Slice, and G. Naylor, eds.). Plenum, New York. (in press).

Klingenberg, C. P., and J. R. Spence. 1993. Heterochrony and allometry: Lessons from the water strider genus *Limnoporus*. Evolution 47:1834–1853.

Klingenberg, C. P., and M. Zimmermann. 1992. Static, ontogenetic, and evolutionary allometry: A multivariate comparison in nine species of water striders. Am. Nat. 140:601–620.

Leamy, L., and W. R. Atchley. 1984. Morphometric integration in the rat (*Rattus* sp.) scapula. J. Zool. Lond. 202:43–56.

Leamy, L., and J. M. Cheverud. 1984. Quantitative genetics and the evolution of ontogeny. II. Genetic and environmental correlations among age-specific characters in randombred house mice. Growth 48:339–353.

Lynch, M. 1988. Path analysis of ontogenetic data. Pages 29–46 *in* Size-structured populations (B. Ebenmann and L. Persson, eds.). Springer-Verlag, Berlin.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

Marcus, L. F. 1990. Traditional morphometrics. Pages 77–122 *in* Proceedings of the Michigan morphometrics workshop (F. J. Rohlf and F. L. Bookstein, eds.). Univ. Mich. Mus. Zool. Spec. Publ. 2. Univ. Michigan, Ann Arbor.

McKinney, M. L., and K. J. McNamara. 1991. Heterochrony: The evolution of ontogeny. Plenum, New York.

Monteiro, L. S., and D. S. Falconer. 1966. Compensatory growth and sexual maturity in mice. Anim. Prod. 8:179–192.

Neuenschwander, B. E. 1991. Common principal components for dependent random vectors. Ph.D. Thesis, University of Bern, Bern, Switzerland.
Pimentel, R. A. 1979. Morphometrics: The multivariate analysis of biological data. Kendall/Hunt, Dubuque, Iowa.
Pitman, E. J. G. 1937. Significance tests which may be applied to samples from any populations. II. The correlation coefficient test. J. R. Stat. Soc. B 4:225–232.
Riska, B., W. R. Atchley, and J. J. Rutledge. 1984. A genetic analysis of targeted growth in mice. Genetics 107:79–101.
Shea, B. T. 1985. Bivariate and multivariate growth allometry: Statistical and biological considerations. J. Zool. Lond. 206:367–390.
Tanaka, A. 1981. Regulation of body size during larval development in the German cockroach, *Blatella germanica*. J. Insect Physiol. 27:587–592.
Tanner, J. M. 1963. Regulation of growth in size in mammals. Nature 199:845–850.
Voss, R. S., and L. F. Marcus. 1992. Morphological evolution in muroid rodents. II. Craniometric factor divergence in seven Neotropical genera, with experimental results from *Zygodontomys*. Evolution 46:1918–1934.
Voss, R. S., L. F. Marcus, and P. Escalante. 1990. Morphological evolution in muroid rodents. I. Conservative patterns of craniometric covariance and their ontogenetic basis in the neotropical genus *Zygodontomys*. Evolution 44:1568–1587.
Wagner, G. P. 1990. A comparative study of morphological integration in *Apis mellifera* (Insecta, Hymenoptera). Z. Zool. Syst. Evolutionsforsch. 28:48–61.
West, T. L., and J. D. Costlow. 1987. Size regulation in larvae of the crustacean *Balanus eburneus* (Cirripedia: Thoracica). Mar. Biol. 96:47–58.
Westfall, P., and S. S. Young. 1993. Resampling-based multiple testing. Wiley, New York.
Zelditch, M. L. 1987. Evaluating models of developmental integration in the laboratory rat using confirmatory factor analysis. Syst. Zool. 36:368–380.
Zelditch, M. L. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. Evolution 42:28–41.

APPENDIX

Principal Component Analysis

Classical PCA deals with observations in a single group, i.e., with a p-variate random vector $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ with covariance matrix $\boldsymbol{\Sigma}$. The PCs, $\mathbf{U} = (U_1, U_2, \ldots, U_p)$, are linear combinations of the original variables, $\mathbf{U} = \mathbf{X}\boldsymbol{\beta}$. This transformation is achieved by the matrix of eigenvectors, $\boldsymbol{\beta}$, which is orthogonal and normalized so that $\boldsymbol{\beta}'\boldsymbol{\beta} = \mathbf{I}_p$.

The covariance matrix of the PCs,

$$\mathrm{Cov}(\mathbf{U}) = \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta} = \boldsymbol{\Lambda},$$

is diagonal, as the PCs are uncorrelated among each other, i.e.,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}.$$

The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the variances of the corresponding PCs (further information can be found in Pimentel [1979], Jolliffe [1986], Flury [1988: chapter 2], Jackson [1991], and Jobson [1992]).

Common Principal Components for Independent Groups

The CPC model for independent groups (Flury, 1988) assumes that all groups share the same eigenvectors. This means that the transformation given by the common matrix of eigenvectors, $\boldsymbol{\beta}$, renders all covariance matrices to diagonal form simultaneously,

$$\boldsymbol{\beta}'\boldsymbol{\Sigma}_i\boldsymbol{\beta} = \boldsymbol{\Lambda}_i, \quad i = 1, \ldots, k,$$

where the $\boldsymbol{\Lambda}_i$ are diagonal matrices (as above) and k is the number of groups.

Common Principal Components for Dependent Random Vectors

The CPC model for dependent random vectors considers kp variables simultaneously, which have a $kp \times kp$ covariance matrix that shows a pattern of $k \times k$ blocks, each of size $p \times p$ (shown here for $k = 2$):

$$\begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

The diagonal blocks $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ are the within-group covariance matrices (as in the previous section), whereas the off-diagonal blocks contain covariances of measurements in different groups ($\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}'$).

The CPC model assumes that the same transformation (using the $p \times p$ orthogonal matrix $\boldsymbol{\beta}$), when applied to all groups, simultaneously renders all blocks diagonal. Therefore, the covariance matrix after transformation to CPCs is

$$\begin{bmatrix} \boldsymbol{\beta}'\boldsymbol{\Sigma}_{11}\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\Sigma}_{12}\boldsymbol{\beta} \\ \boldsymbol{\beta}'\boldsymbol{\Sigma}_{21}\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\Sigma}_{22}\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} \end{bmatrix},$$

where all $\boldsymbol{\Lambda}_{ij}$ are diagonal. This model is discussed in detail by Neuenschwander (1991) and Flury and Neuenschwander (1995a); algorithms for estimating CPCs were presented by Flury and Neuenschwander (1995b).

Number of Parameters

The number of parameters in the unconstrained model is

$$pk(pk + 1) / 2.$$

Under the CPC model for dependent random vectors, this number is

$$p(p - 1) / 2 + pk(k + 1) / 2,$$

where the first term accounts for the CPC coefficients and the second term for the within-group variances of the CPCs and their covariances across groups.

As p and k increase, the reduction in parameters under the CPC model becomes very substantial.

Estimation of CPCs

In a sample, the $kp \times kp$ covariance matrix $\mathbf{S}$ is patterned as explained above for $\boldsymbol{\Sigma}$ (again illustrated for $k = 2$), i.e.,

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}.$$

Then we search for an orthogonal $p \times p$ matrix $\mathbf{B}$ (normalized so that $\mathbf{B}'\mathbf{B} = \mathbf{I}_p$) that simultaneously renders the four blocks of $\mathbf{F}$ as closely to diagonal as possible, where

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{B}'\mathbf{S}_{11}\mathbf{B} & \mathbf{B}'\mathbf{S}_{12}\mathbf{B} \\ \mathbf{B}'\mathbf{S}_{21}\mathbf{B} & \mathbf{B}'\mathbf{S}_{22}\mathbf{B} \end{bmatrix}$$

is the covariance matrix of the transformed variables.

The measure of deviation from simultaneous diagonality is

$$e = \frac{\det \begin{bmatrix} \operatorname{diag} \mathbf{F}_{11} & \operatorname{diag} \mathbf{F}_{12} \\ \operatorname{diag} \mathbf{F}_{21} & \operatorname{diag} \mathbf{F}_{22} \end{bmatrix}}{\det \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}},$$

where "det" is the determinant of a matrix and the "diag" operator sets the off-diagonal elements of a matrix to zero. It can be shown that e is a minimum if all $\mathbf{F}_{ij}$ are diagonal. The *FG*+ algorithm (Flury and Neuenschwander, 1995b) is designed to find an orthogonal matrix $\mathbf{B}$ that minimizes this measure (for further discussion, see Neuenschwander [1991] and Flury and Neuenschwander [1995a, 1995b]).

Individual Variation of Ontogenies: A Longitudinal Study of Growth and Timing*

INTRODUCTION

The evolution of ontogeny has attracted much attention in recent years (Gould 1977; McKinney and McNamara 1991; Hall 1992). Interest in the evolution of organismic form motivated part of this research, because the diversity of morphological structures is the outcome of variation in growth and development. On the other hand, life history studies include growth in size and the schedules of transitions between developmental stages as important elements of the relationships between organisms and their environment (Roff 1992; Stearns 1992).

Some studies in this field have compared the ontogenies of several taxa, viewing differences as the results of past evolutionary change (e.g., Creighton and Strauss 1986; Strauss 1990; Klingenberg and Spence 1993). Others have dealt with the patterns of variation of ontogenetic processes within populations that provide the potential for continuing microevolution (e.g., Cheverud et al. 1983; Lynch 1988; Kirkpatrick and Lofsvold 1989; Atchley and Hall 1991). In such analyses of ontogenetic variation, the lack of variability can be as important as its presence, because it constitutes a developmental constraint on the future evolution of the traits under consideration (Maynard Smith et al. 1985; Gould 1989; Kirkpatrick and Lofsvold 1992; Björklund 1993).

Developmental processes produce morphological variation and constraints and thus affect evolutionary processes in two principal ways. First, they determine the growth curves, that is, the functions that relate age or developmental stage to morphological and physiological traits. As a consequence, they influence the extent to which values of the same trait at different ages can vary independently (Cheverud et al. 1983; Kirkpatrick and Lofsvold 1992; Björklund 1993). Second, because developmental processes often affect several traits simultaneously (Riska 1986), multivariate patterns of covariation among traits at a given age reflect these processes (e.g., Cheverud 1982a, 1995; Zelditch 1987; Cowley and Atchley 1990; Atchley et al. 1992; Paulsen and Nijhout 1993), and can in turn affect the potential for evolutionary change (Lande 1979; Cheverud 1984). These two aspects have been integrated in theoretical syntheses (Atchley and Hall 1991; Cowley and Atchley 1992; Atchley et al. 1994), but there are few detailed empirical studies that jointly consider covariation among traits and across developmental stages (but see Cheverud et al. 1983; Björklund 1993). Such studies have been limited because statistical techniques specifically designed for longitudinal studies with multiple measurements have not been available. Furthermore, the existing studies of ontogenetic variation and constraints have been based on analyses of age-specific measurements, but few have examined the variation in growth increments (but see Riska et al. 1984).

---

* A version of this chapter is submitted for publication. Klingenberg. 1996. Evolution.

Growth processes not only produce variation in morphometric traits, but they can also eliminate it by compensatory growth, so that all individuals converge toward a "target" size as adults (Tanner 1963), irrespective of differences in their earlier growth history. Differences in growth may be due to starvation (Wilson and Osbourn 1960; Blum et al. 1985) or individual variation apparent even under controlled laboratory conditions (Monteiro and Falconer 1966; Atchley 1984; Riska et al. 1984). Although vertebrates have been studied in the most detail, growth regulation has also been reported from insects (Tanaka 1981; Bryant and Simpson 1984), crustaceans (Hartnoll and Dalley 1981; West and Costlow 1987) and echinoderms (Ettensohn and Malinda 1993). Whereas some regulation of growth occurs through hormones (Tanner 1963; Blum et al. 1985; Shea et al. 1990) and thus affects multiple measurements simultaneously, leading to tight overall integration, there is also ample evidence from a variety of studies that organs independently control their final size to a considerable degree (Bryant and Simpson 1984). Therefore, although growth regulation potentially is a major determinant of ontogenetic variation, it is not possible to predict what specific effects these processes have on the patterns of covariation among morphometric traits.

Ontogenetic variability is not limited to morphometric characters, but the timing of developmental events also can vary, which may lead to evolution by heterochrony (McKinney and McNamara 1991). To study this variation within and between populations, it is important to find appropriate standards for comparison. Yet studies of organisms with continuous growth most commonly are forced to ignore this variation, because they are usually based on measurements at fixed ages (e.g., Cheverud et al. 1983; Riska et al. 1984), even in comparisons between taxa (e.g., Creighton and Strauss 1986; Björklund 1993; Wayne and Ruff 1993). A possible alternative is standardization of age relative to an ontogenetic event, such as the peak velocity of growth in height for human adolescents (Cameron et al. 1994), but this may be very difficult if the event is gradual and if the data have high temporal resolution. To avoid such ambiguities, an ideal study system should have a fixed number of discrete developmental stages, a condition that is met by many hemimetabolous insects.

Water striders (Heteroptera: Gerridae) are perfect study organisms to address these questions because they can be reared individually in the laboratory, measurements of the cuticles cast during molting provide an accurate record of growth, and each molt is a distinct developmental event. Here I report the results of a longitudinal growth study in *Limnoporus canaliculatus* (Say). This study thus complements an earlier comparison of ontogenies among the six species of the genus *Limnoporus*, which revealed a considerable degree of interspecific variation in timing and extent of growth (Klingenberg and Spence 1993). I carried out a joint analysis of covariation among measurements within and across developmental stages, using a new statistical model specifically designed for such studies (Flury and Neuenschwander 1995a; Klingenberg et al. 1996). Based on these results, I compare analyses of variation and constraint in growth increments to those in age-specific measurements, and I examine growth regulation and the relation between instar durations and growth in size.

MATERIALS AND METHODS

This study is based on longitudinal data from all five larval instars (L1–L5) and adults of the water strider *Limnoporus canaliculatus*. Exuviae collected from bugs reared individually make it possible to obtain measurements from single individuals in all growth stages without manipulating the delicate larvae.

Laboratory Culture and Measurements

The water striders used in this study were the offspring of a sample of overwintered adults collected in Morris County, New Jersey, on May 1, 1992. The entire laboratory rearing experiment was carried out in the same climate-controlled room (20° C, photoperiod 16L:8D). Adults were kept as a mass culture, provided with Styrofoam strips for oviposition, and fed ad libitum with frozen flesh flies, *Neobelliera bullata* (Parker). Styrofoam strips were replaced regularly, and those with eggs in advanced stages of development were checked for hatched larvae at intervals of approximately 12 hours.

Hatchlings were transferred into individual rearing containers (diameter 11.5 cm, height 8 cm), each with about 1 cm of water and a small piece of Styrofoam floating on the surface. Each larva was fed a frozen flesh fly daily; this is an ad libitum regime, as the weight of a fly far exceeds that of even an adult water strider. Larvae were checked for molts twice daily; the data for instar durations therefore have a resolution of approximately 12 hours. After each molt, the cast exuvia was collected and subsequently stored in 70% ethanol. After the final molt, when the new cuticle had hardened, adults were killed by deep-freezing and later stored in 70% ethanol.

For this study I analyzed measurements of the lengths of the femora and tibiae of the middle and hind legs. The cuticle of the legs is rigidly sclerotized; therefore, shrinking or other effects of preservation can be ruled out. Because antennal segments tended to telescope into one another in exuviae, their lengths, included previously in cross-sectional studies of growth in water striders (Klingenberg and Zimmermann 1992a; Klingenberg and Spence 1993), could not be measured reliably and were therefore not used in this study. Measurements were made with a video system attached to a dissecting microscope.

The data analyzed in this study are means of left and right body sides if both sides could be measured; if the value from one side was missing, the value measured on the other side was included. The combined variability from asymmetry and measuring error was small relative to the variation among individuals. The data set includes only those individuals for which all four variables could be measured on at least one body side in all five larval instars and the adult stage. I checked data for outliers, and reexamined individuals with extreme values. Bugs with deformities, mostly because of abnormal molting, were excluded. Wing polymorphism did not show an influence on the results of this study. Most of the individually reared water striders were wingless: of the 89 females with complete data, only one was winged, and of the 70 males, five were winged (two of them brachypterous), although the mass culture in the same room produced a higher proportion of winged bugs. As preliminary univariate and multivariate analyses showed that winged individuals did not differ from the wingless ones either in morphometric traits or in instar durations, I included all bugs in this study regardless of wing morph.

All the morphometric variables, but not the instar durations, were transformed to natural logarithms before the analyses.

Statistical Analyses

Longitudinal growth studies with multiple measurements are complex because there are correlations both within and across ontogenetic stages. If there are p measurements for each individual at k growth stages, the overall covariance matrix $\mathbf{S}$ has a pattern of k^2 blocks, each of dimension $p \times p$. The block $\mathbf{S}_{ij}$, in the i-th row and the j-th column of $\mathbf{S}$, is a $p \times p$ matrix that contains the covariances of the p measurements in stage i with those in stage j ($\mathbf{S}_{ij}' = \mathbf{S}_{ji}$). The blocks along the diagonal, $\mathbf{S}_{ii}$, are the within-stage covariance matrices, as they have been used traditionally in both longitudinal and cross-sectional studies (e.g., Cuzin-Roudy 1975; Zelditch and Carmichael 1989; Klingenberg and Zimmermann 1992a). Conversely, analyses of the covariation among stages in one measurement at a time (e.g., Cheverud et al. 1983; Kirkpatrick and Lofsvold 1989, 1992; Björklund 1993), say the h-th variable, consider only the h-th position along the diagonal in each of the blocks $\mathbf{S}_{ij}$.

Common Principal Component Model.—The basis of the statistical model I use here is the observation that patterns of variation among characters are often similar within several ontogenetic stages (e.g., Cuzin-Roudy 1975; Zelditch and Carmichael 1989; Klingenberg and Zimmermann 1992a), but the model extends this similarity to all blocks of the covariance matrix. More specifically, it assumes that the different stages share the same principal components, and that these are mutually uncorrelated not only within, but also across developmental stages (Klingenberg et al. 1996). This model, common principal components (CPCs) for dependent random vectors (Neuenschwander 1991; Flury and Neuenschwander 1995a), is an extension of the CPC model for independent groups (Flury 1988), which has been used in a number of morphometric studies (Airoldi and Flury 1988; Klingenberg and Zimmermann 1992a,b; Klingenberg and Spence 1993; Klingenberg 1996).

For longitudinal data, this CPC model can reduce a very complex analysis to a number of separate, simpler ones. As the CPCs are uncorrelated within and across stages, each block of the patterned matrix of CPC scores is diagonal; the diagonal elements can be used to study covariation of CPCs among stages. Unlike the original measurements, where separate analyses for each trait of the covariation across stages ignore the correlations between variables, each CPC can be analyzed in isolation without any loss of information because they are uncorrelated with each other. Like conventional principal component analysis, this CPC model can be useful as a tool for data reduction, because some components often account for only a minor fraction of the total variation and may be ignored. The fit of the model can be assessed either by tests based on asymptotic theory (Neuenschwander 1991) or by permutation tests (Klingenberg et al. 1996).

Klingenberg et al. (1996) have demonstrated the use of the CPC model for the data set of female *L. canaliculatus* and have discussed it in some detail; here I also apply it to the males, but I mention the statistical aspects only briefly. To estimate CPCs, I used a version of the orthogonal FG^+ algorithm (Flury and Neuenschwander 1995b) written in the SAS/IML language (this

routine is available through the Internet: file://life.bio.sunysb.edu/morphmet/dcpc.exe.ibmpc). The CPCs were ordered by the average proportion of total variance within instars and covariance among instars for which they accounted (for details, see Klingenberg et al. 1996).

I tested the fit of the CPC model to the data with a permutation test. The key assumption of the model is that different CPCs are uncorrelated within and between instars, whereas corresponding CPCs can be correlated across instars (e.g., the CPC1 in the L1 can be correlated with the CPC1 in the L2, but not with the CPC2 in any instar). This can be used as the null hypothesis in a permutation test (Pitman 1937; Manly 1991; Good 1994). For each sex, I ran 1,000 random permutations: the CPC scores were reshuffled separately for the CPC2–CPC4, but keeping all instars together, thus leaving unchanged the correlations among instars for each CPC. For each of the randomized data sets, I computed the CPCs and three different test statistics. The e statistic (Klingenberg et al. 1996) is a measure of overall deviation from the CPC model (for a detailed discussion, see Neuenschwander 1991; Flury and Neuenschwander 1995a,b). I used two additional statistics, the maximum absolute correlation and covariance between different CPCs, because they can pinpoint the CPCs and instars where deviations occur (note that ordinary significance tests do not apply here, because this is the highest absolute value of the 216 correlations or covariances between different CPCs). The null distributions of the test statistics from the permutation runs were then compared to the values from the original data.

To compute standard errors, I used the bootstrap method (Efron and Tibshirani 1993), with 250 iterations. A preliminary bootstrap analysis produced inflated standard errors because of changes in the ordering of CPCs and in the signs of their coefficients (see also Jackson 1993). To avoid this problem, I used a different rule to order the CPCs for the bootstrap routine, and assigned each bootstrap CPC to that CPC of the original sample to which it was most similar, as judged by the magnitude of their inner product (this is equivalent to the use of angles between CPCs); these assignments were unambiguous in all 250 bootstrap runs for each sex. To prevent arbitrary changes in signs of CPC coefficients, the signs of all coefficients of the bootstrap CPC were reversed if the inner product of the original and corresponding bootstrap CPCs was negative.

Ontogenetic Allometry.—To take advantage of the information contained in the longitudinal data, I computed patterns of ontogenetic allometry in a manner slightly different from cross-sectional studies (e.g., Klingenberg and Zimmermann 1992a). Instead of a PCA of data pooled over individuals and stages, here I used a MANOVA of individuals and instars to separate static (individual) from ontogenetic variation. Therefore, a PCA of the between-instar matrix of sums of squares and cross-products can be used to analyze ontogenetic variation, and the resulting PC1 is a vector of ontogenetic allometry. To estimate standard errors, I used the bootstrap procedure, with 250 resampling iterations (random resampling among individuals, i.e., keeping the measurements from all instars together for each bug). A separate analysis was run for each sex.

Analyses of Variation and Constraints.—To study patterns of variation and identify possible constraints on the dynamics of growth in overall size, I used conventional principal component analyses (PCAs) of the covariance matrices

of CPC1 scores in the six instars, of the increments in these "size" scores, and of instar durations. To estimate standard errors for these PCAs, I used the bootstrap with 250 resampling iterations. Inflated standard errors due to changes in the ordering of PCs and sign reversals of PC coefficients were a problem in some of these analyses, and as for the CPCs, I assigned the bootstrap PCs to the most similar PC in the original sample (i.e., the one with which it had the highest absolute inner product), and changed signs if the inner product was negative. Similar analyses were conducted for the CPC2 scores and increments.

For analyzing growth and its regulation, I derived a log-size variable (Mosimann 1970) from the CPC1 by rescaling its coefficients so that they summed up to unity; its antilogarithm therefore scaled as a linear dimension (Klingenberg and Zimmermann 1992b; Klingenberg and Spence 1993). I computed growth ratios as the antilogarithm of the difference in the log-size scores between successive instars. More intuitively, this measure of growth can be interpreted as the postmolt/premolt ratio for overall size. The geometric mean of these ratios in a sample is the geometric-mean growth ratio, which can also be obtained from cross-sectional studies (Klingenberg and Zimmermann 1992b; Klingenberg and Spence 1993). As a measure of relative size within instars, I used the ratio of the individual's size to the geometric mean size in that instar, computed as the antilogarithm of the difference of each individual's log-size score from the instar mean score.

Correlations among relative size, postmolt/premolt ratios, and instar durations are Pearson product-moment correlations. I tested them against the null hypothesis of independence with two-tailed permutation tests (Pitman 1937), each with 10,000 random permutations (see also Manly 1991; Good 1994). I present correlations with their original P-values, but to determine statistical significance, I use the sequential Bonferroni adjustment (Rice 1989) to control for experimentwise error rate within each sex (table-wide $\alpha = 0.05$).

RESULTS

Static Variation: Common Principal Components

The analyses in both sexes produced similar results (table 6-1). The CPC1 is a size axis, whose coefficients are all positive and of similar magnitude. The CPC2 shows a contrast of the middle and hind femora against the hind tibia. The CPC3 opposes the middle tibia to the hind femur and, to a lesser extent, to the hind tibia. Finally, the CPC4 contrasts the middle femur to the middle tibia and, more strongly, to the hind femur. The close congruence of results between sexes is even more apparent from vector correlations between pairs of CPCs, which all exceed 0.99; the corresponding angles between CPC axes range from 5.7° to 7.8°. The CPC estimates are fairly stable, as indicated by their standard errors. The standard errors obtained with the bootstrap method are fairly similar (differences < 0.03) to those from a jackknife analysis (for females only; Klingenberg et al. 1996).

Table 6-1. Common principal component coefficients and their bootstrap standard errors (in parentheses). Abbreviations of morphometric variables: MF, middle femur; MT, middle tibia; HF, hind femur; HT, hind tibia.

Variable	CPC1	CPC2	CPC3	CPC4
Females				
MF	0.471	0.436	–0.059	0.765
	(0.019)	(0.038)	(0.115)	(0.029)
MT	0.414	0.073	0.878	–0.229
	(0.024)	(0.049)	(0.062)	(0.136)
HF	0.491	0.465	–0.427	–0.600
	(0.026)	(0.042)	(0.108)	(0.084)
HT	0.605	–0.767	–0.209	0.048
	(0.030)	(0.031)	(0.052)	(0.054)
Males				
MF	0.491	0.462	0.003	0.738
	(0.031)	(0.038)	(0.082)	(0.022)
MT	0.388	0.059	0.870	–0.299
	(0.016)	(0.054)	(0.035)	(0.097)
HF	0.560	0.343	–0.475	–0.586
	(0.026)	(0.067)	(0.077)	(0.056)
HT	0.542	–0.816	–0.135	0.150
	(0.055)	(0.040)	(0.060)	(0.035)

All four CPCs account for fairly constant proportions of variation throughout ontogeny (table 6-2). The CPC1 takes up more than 75% of the total variation in all instars and both sexes. A moderate amount of variation, about 10–16%, is associated with the CPC2, whereas the other two CPCs show substantially less variability. These values are close to the proportions of total variance for which the PCs accounted in separate PCAs in each instar and sex (for the females, this comparison is presented in Klingenberg et al. 1996).

Table 6-2. Percentages of total variance for which the CPCs account within each instar, and their bootstrap standard errors (in parentheses).

Instar	CPC1	CPC2	CPC3	CPC4
Females				
L1	75.9	13.8	6.7	3.6
	(3.8)	(2.6)	(1.5)	(0.8)
L2	79.7	11.9	5.3	3.1
	(3.8)	(2.6)	(1.2)	(0.7)
L3	76.8	14.0	5.4	3.9
	(3.9)	(2.7)	(1.3)	(0.9)
L4	79.7	13.2	5.1	1.9
	(3.5)	(2.6)	(1.1)	(0.5)
L5	79.8	13.2	5.2	1.8
	(3.3)	(2.5)	(1.2)	(0.5)
Ad.	78.2	14.4	5.8	1.5
	(3.6)	(2.8)	(1.4)	(0.4)
Males				
L1	76.3	16.2	4.0	3.6
	(3.6)	(3.1)	(0.7)	(0.9)
L2	79.0	13.2	5.0	2.8
	(2.6)	(2.3)	(1.1)	(0.5)
L3	81.2	10.5	4.8	3.5
	(3.0)	(2.2)	(1.3)	(0.7)
L4	80.5	11.7	5.2	2.6
	(3.6)	(2.4)	(1.4)	(0.6)
L5	81.9	12.0	4.5	1.6
	(4.4)	(4.1)	(1.4)	(0.3)
Ad.	79.8	13.7	4.6	1.9
	(3.6)	(3.0)	(1.3)	(0.6)

The permutation tests reveal some deviations from the CPC model for both sexes. In females, all three test statistics are significant, but the deviations concern exclusively the CPC3, which accounts only for a small proportion of variance, and will not be considered here (table 6-2; for details, see Klingenberg et al. 1996). In males, the e statistic indicates a significant deviation from the CPC model overall ($e = 405.9$; $P < 0.001$), and the maximum absolute covariance between different CPCs (value 8.48; $P = 0.01$) attributes it to the rela-

tively large covariance between CPC1 and CPC2 in adults, which both have large variances (none of the other combinations of CPCs have similarly large covariances). Nevertheless, the maximum absolute correlation does not indicate a statistically significant deviation from independence (maximal $|r| = 0.36$, CPC1 in L1 with CPC3 in L2; $P = 0.19$). In both sexes the first two CPCs, which account for about 90% of the variation in each instar, can be considered uncorrelated within and between instars, despite the misfit of the model overall, and therefore can be studied separately in the subsequent analyses.

Relations between Static and Ontogenetic Allometry

In analyses of ontogenetic allometry, the PC1s account for the overwhelming majority of variation, 99.7% in females and 99.8% in males, and thus indicate that the model of simple allometry fits the data very well. The patterns of ontogenetic and static allometry are similar (fig. 6-1), as indicated by angles of 4.5° and 6.5° between ontogenetic PC1 and static CPC1 in males and females. respectively (vector correlations > 0.99). The angles between the ontogenetic PC1 and the static PC1s for each instar are narrowest in the L2 or L3 instars, and increase toward the adult stage.

Individual Variation in Growth

To assess the patterns of instar-specific variability in overall size, I performed a PCA of the CPC1 scores in each instar. The first PC alone accounts for about three quarters of the total size variation in both sexes (fig. 6-2). This PC1 is an axis summarizing variation in general growth performance: all instars have positive coefficients (fig. 6-2), indicating that individuals tend to be either relatively large or relatively small in all instars. The PC1 coefficients gradually increase from instar to instar, manifesting variation in slope that is associated with the height of growth curves, i.e., larger bugs also tend to have the steeper growth curves. The PC2, which takes up about 17% of the total

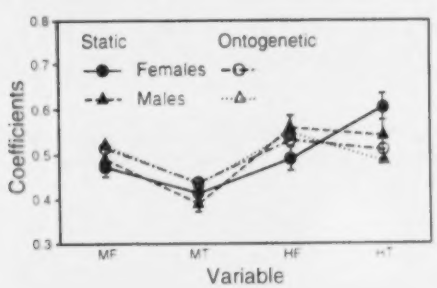


Fig. 6-1. Comparison of static and ontogenetic allometry. Joint patterns of static allometry in all six instars were estimated by the CPC1 coefficients, and ontogenetic allometry by the PC1s of the between-instar matrices from MANOVA. Error bars indicate bootstrapped standard errors (250 iterations); note that those for the ontogenetic PC1s are extremely small (all ≤ 0.001) and entirely covered by the symbols.

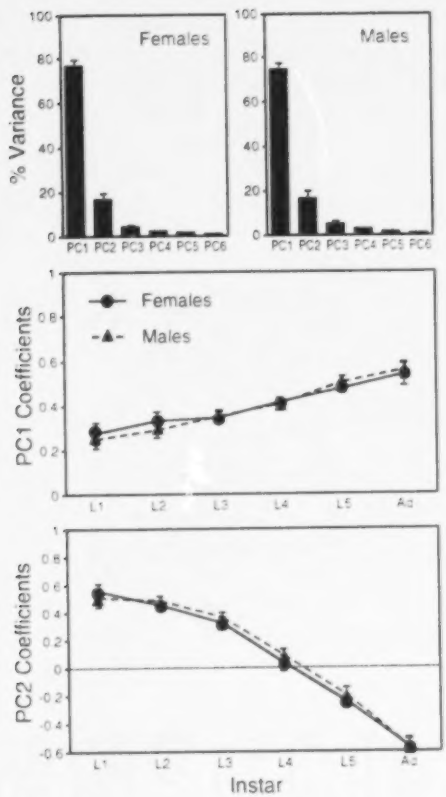


Fig. 6-2. Variation and constraint in instar-specific size. The PCA used the covariance matrix of individual CPC1 scores in all six instars. The top panels show the eigenvalues, expressed as percentages of total variance; note that the PC1 takes up most of the variation. The middle and bottom panels show the coefficients of each instar on the PC1 and PC2, respectively. The PC1 has positive coefficients for sizes in all instars and shows variation in overall growth, whereas PC2 contrasts size in early versus late instars. Error bars indicate the bootstrapped standard errors of the estimates; some of the standard errors are so small that the error bars are too short to be seen or the symbols entirely cover them.

variance in both sexes, is the only other PC accounting for a relatively large proportion of variation. It contrasts the overall size scores in early against those in late instars in a graded series (fig. 6-2), showing that growth curves vary in slope, but not in shape, as the profile of PC2 coefficients is fairly straight. The PC2 thus features variation only in the slope of growth curves; it can be visualized as a movement where the entire growth curve oscillates like a "seesaw" pivoting about the fairly constant size in the L4 instar, but does not "bend". The remaining PCs only account for small proportions of variance.

Given the dominance of the PC1 in the preceding analysis, it is rather surprising that the growth increments are not strongly correlated between instars

(fig. 6-3). In females, correlations between postmolt/premolt ratios of the overall size variable range from –0.22 (L2–L5) to 0.47 (L3–L4; $P = 0.0001$); the latter and the correlation between L4 and L5 increments ($r = 0.32$; $P = 0.002$) are the only ones significantly different from zero after sequential Bonferroni adjustment for the 10 pairwise correlations. In males, correlations range from –0.09 (L1–L5) to 0.34 (L3–L4; $P = 0.004$, and thus significant after adjustment; all others nonsignificant).

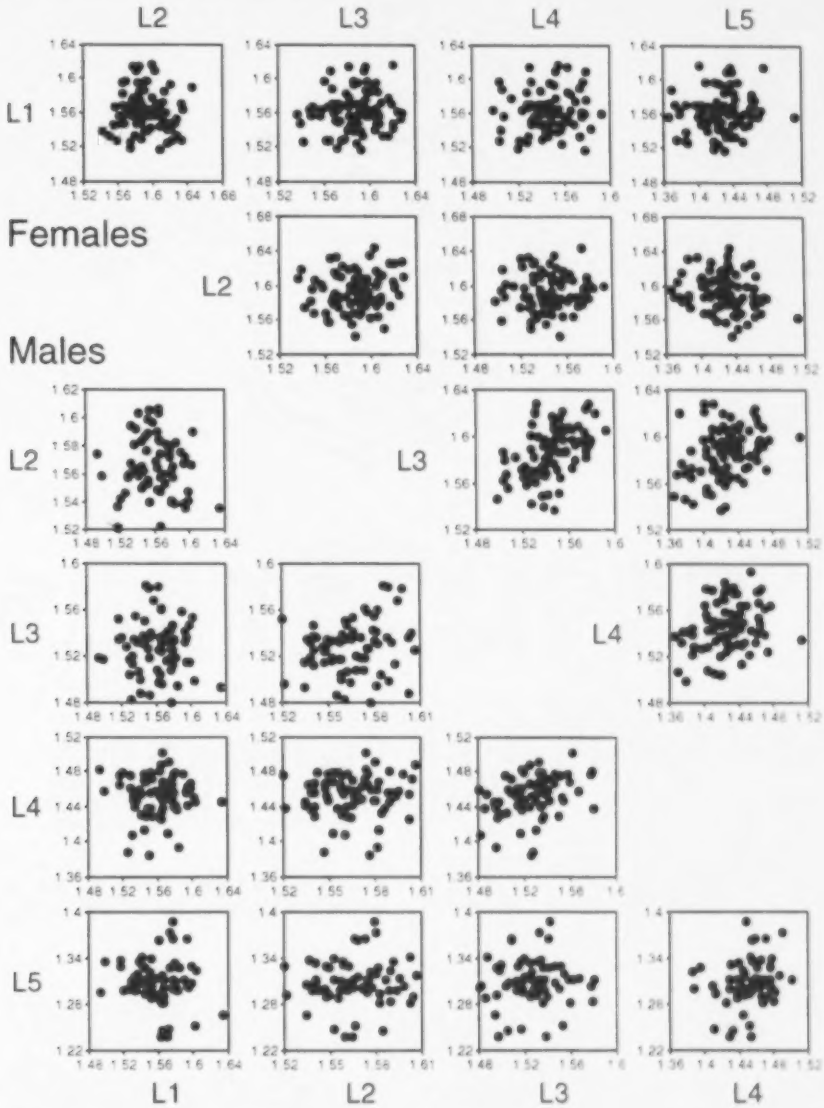


Fig. 6-3. Relations between growth increments of individual water striders in different instars. Plotted values are the postmolt/premolt ratios of a multivariate size measure that scales as a linear dimension (see text for details).

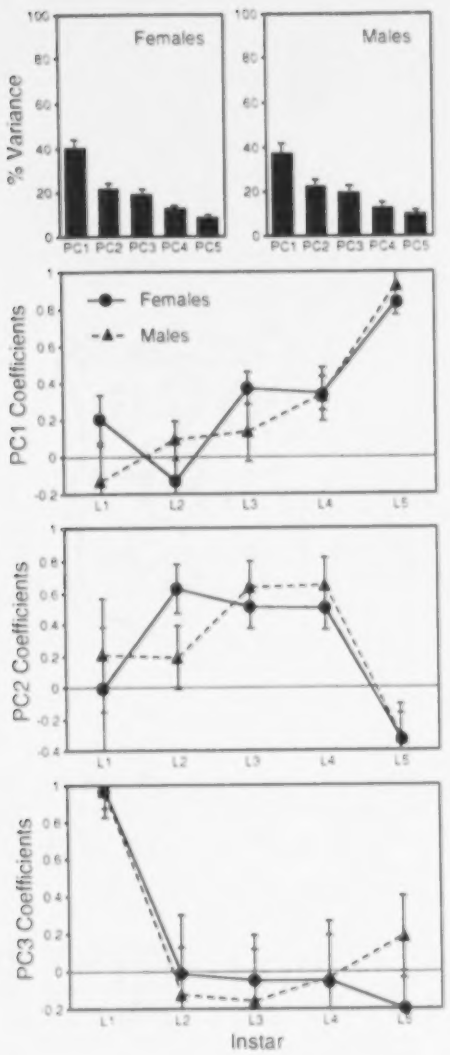


Fig. 6-4. Patterns of variation in size increments. The PCA was based on the covariance matrix of differences in CPC1 scores between successive instars. The eigenvalues, shown as percentages of the total variation (top panels), indicate that the PC1 is less dominant in this analysis of increments than in the corresponding analysis for instar-specific CPC1 scores (fig. 6-2). The lower three panels show the coefficients of each instar on the PC1, PC2, and PC3; none of these is an axis featuring variation in growth performance throughout ontogeny. Error bars are bootstrapped standard errors of the respective estimates.

The PCAs of increments in CPC1 scores reflect this weak covariation, as the PC1s of increments account for much smaller fractions of the total variation (fig. 6-4) than in the analyses of instar-specific size (fig. 6-2). Moreover, the larger standard errors and the incomplete congruence of results between the

sexes indicate that patterns are less well-defined. In the analysis of increments, none of the PCs is an axis of variation in growth performance in all instars jointly (fig. 6-4). Instead, the PC1s almost exclusively feature the variability in late growth, as they have much larger coefficients for the L5 than for earlier instars (fig. 6-4). The PC2s emphasize increments in the L3 and L4 (and L2 in females), and contrast this to L5 growth, whereas the size increase in the first molt alone dominates the PC3s. Although the remaining two PCs account for smaller portions of variation, both are associated with appreciable variability. Therefore, this PCA provides no evidence of any substantial constraints on growth increments, in contrast to the analyses of instar-specific size, where variation is largely concentrated in just two dimensions.

Postmolt/premolt ratios also are at most moderately correlated with relative size in the previous instar (fig. 6-5). In the L1–L3 instars, correlations between relative size and growth ratios are negative, and thus indicate compensatory growth (females: L1 $r = -0.25$, $P = 0.015$, marginally nonsignificant after sequential Bonferroni adjustment; L2 $r = -0.28$, $P = 0.007$; males: L1 $r = -0.33$, $P = 0.004$; the latter two correlations remain significant after adjustment for five tests in each sex). Yet, because these correlations are rather weak, compensatory growth is just strong enough to maintain the variance of the CPC1 approximately constant in the L1 and L2 instars (in females also the L3). In the L3, compensatory growth eliminates less variation than the new growth increments generate, and therefore, morphometric variance increases from the L3 to the L4 instar. In the final two molts, the correlations between size in the previous instar and growth are positive, although not statistically significant, and the variance of the CPC1s increases by 36– 51% in each molt.

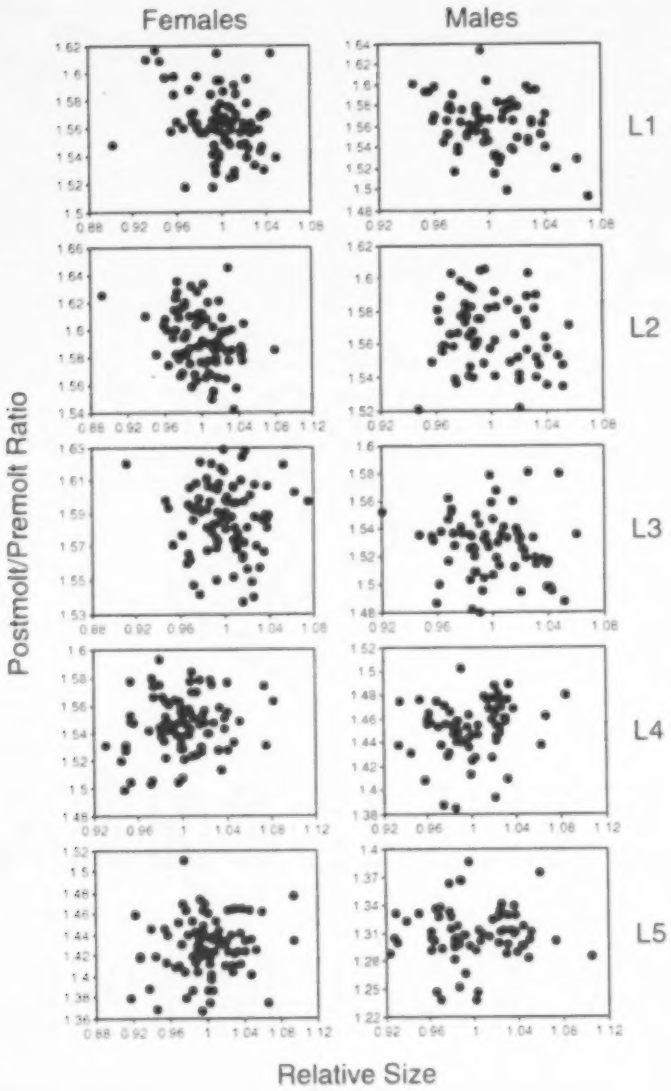


Fig. 6-5. Relations between relative size in an instar and the size increment in the following molt. Relative size is the ratio between an individual's score for the multivariate size variable and the geometric mean for this variable in that instar and sex, and the postmolt/premolt ratio is the ratio between the size scores in this and the following instar. There are some negative correlations of size and growth ratio in early instars, indicating compensatory growth, but not in later instars.

To examine whether the patterns of ontogenetic variation in size-independent variables are the same as for the "overall size" component, I performed

PCAs of the scores and increments of the CPC2. The results are similar to the corresponding analyses for the CPC1. For the CPC2 scores, the PC1s account for 61% and 58%, and the PC2s for 22% and 21% of the total variance in females and males, respectively. The PC1 coefficients are all positive and increase in magnitude from early to late instars, and thus indicate that bugs tend to have either high or low CPC2 scores in all instars, whereas the PC2 coefficients gradually decline from positive to negative values, indicating "seesaw-like" variation of CPC2 growth trajectories. In the analyses of CPC2 increments, like in those for the CPC1, the PC1s account for less of the total variation (39% and 53% in females and males, respectively), and each PC mainly features the increment during one instar, contrasting it with CPC2 changes in other instars. Correlations between CPC2 scores in one instar and the differences to the next instar show that there clearly is growth regulation for this size-free component of variation (with correlations as strong as $r = -0.55$). In females, these correlations are significantly negative in the L1, L2, and the L3, whereas in males the L1, L2, and L5 instars have significant negative correlations (all after sequential Bonferroni correction).

Instar Durations

Unlike the size increments, instar durations are clearly correlated among instars, and the range of instar durations tends to increase from younger to older stages (fig. 6-6). Correlations range from 0.47 (L1–L4) to 0.85 (L2–L4) in females and from 0.38 (L1–L2) to 0.88 (L3–L4) in males. All these correlations are highly significant, and remain so even after Bonferroni correction for the 10 tests, as none of the permutation runs matches the observed values ($P < 0.0001$), except for three correlations in males (P-values between 0.0001 and 0.0015, all significant after sequential Bonferroni adjustment). Altogether, these correlations show that developmental rates of individuals vary consistently in all instars, in contrast to the results found for the size increments (above).

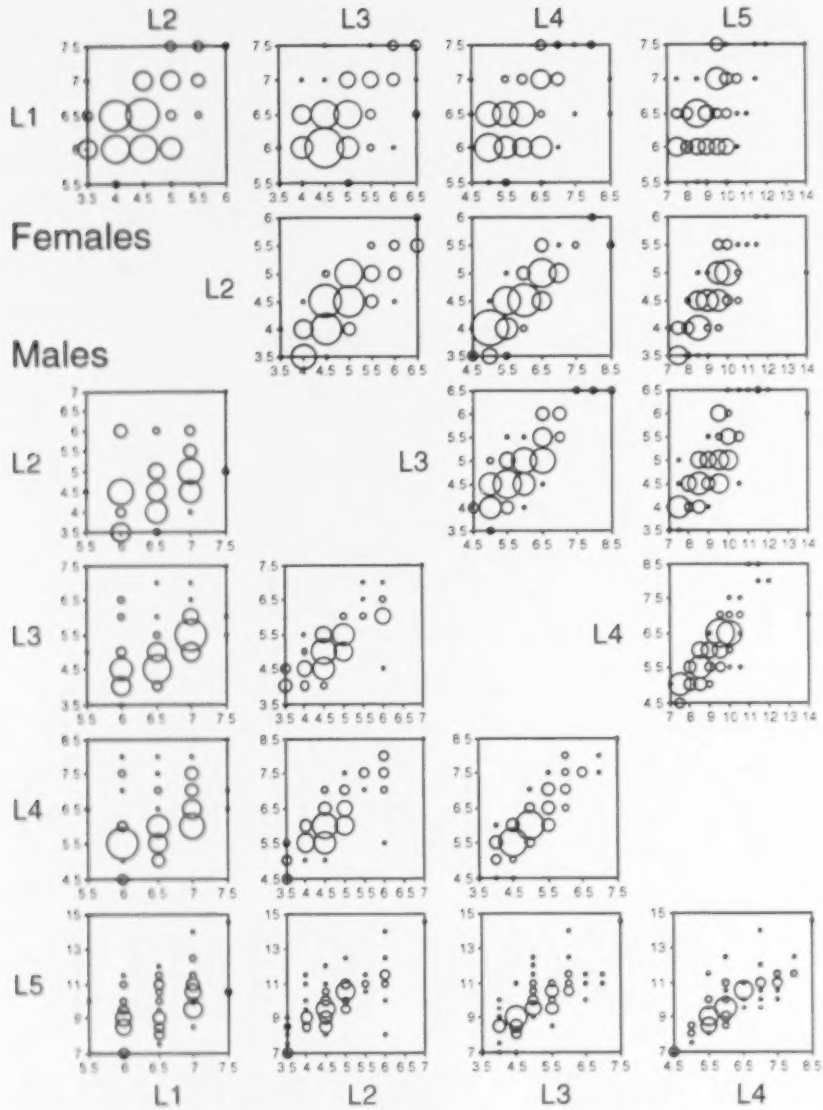


Fig. 6-6. Relations between the durations of different instars. As checks were made at intervals of about 12 hours, the temporal resolution of the data is relatively coarse, and there is extensive overlap of data points. In this graph, therefore, the diameter of the "bubbles" is proportional to the number of individuals with a particular combination of instar durations. Note the increasing variation in instar duration and the successively stronger correlations between instars.

In accordance with these correlations, the PC1 of the covariance matrix of instar durations takes up almost all the total variance (fig. 6-7). All instars have

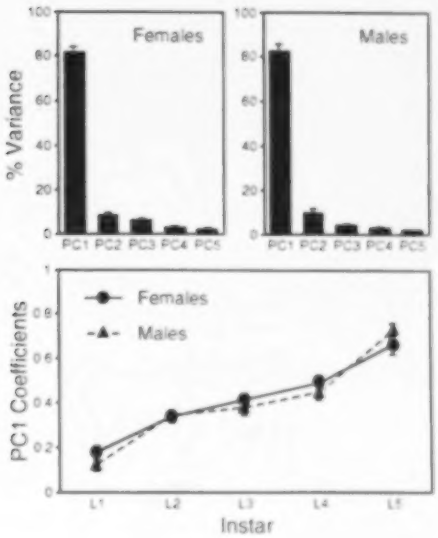


Fig. 6-7. Patterns of variation in instar durations. The PCA used the covariance matrix of untransformed instar durations. The top panels show the eigenvalues, expressed as percentages of total variance; note the strong dominance of the PC1. The bottom panel shows the coefficients of each instar on the PC1. Error bars indicate the bootstrapped standard errors of the estimates; for a few of the PC1 coefficient values, the symbols entirely cover the error bars.

positive PC1 coefficients, which gradually increase from the L1 to the L5, reflecting the larger variation in later instars. The results for both sexes are very similar, and the small standard errors indicate that they are statistically stable.

A PCA of cumulative development time, that is, the ages at the five molts, shows an even stronger dominance of the PC1 (fig. 6-8). As in the preceding analysis, the PC1 coefficients increase from instar to instar, but this increase is steadier and somewhat stronger here.

Instar durations are only weakly correlated with the size increments in the same instar (fig. 6-9). Correlation coefficients are negative or very close to zero, and only those in the L5 instars (and L3 in males) retain statistical significance after sequential Bonferroni adjustment (females, L5, $r = -0.34$, $P = 0.0006$; males, L3, $r = -0.39$, $P = 0.0006$; L5, $r = -0.32$, $P = 0.006$). Yet these correlations are not significant for growth from hatching to the final molt, as relative size of adults and total development time are uncorrelated (females $r = -0.14$, $P = 0.18$; males $r = -0.14$, $P = 0.24$).

Discussion

This study integrates approaches that traditionally have been used separately to study variation in ontogeny. Measurements of multiple traits in each instar provide the basis for analyzing allometry and morphological integration, and the longitudinal nature of the data enabled me to examine growth regulation

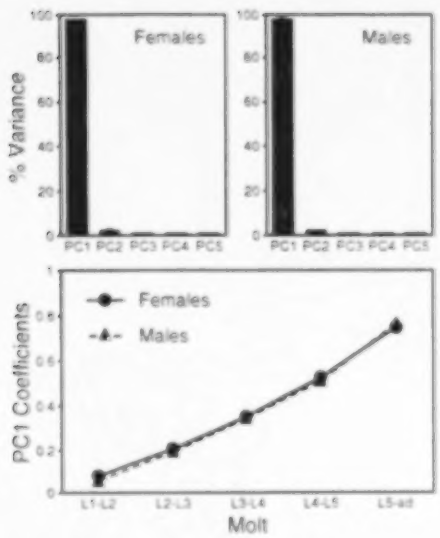


Fig. 6-8. Patterns of variation in cumulative development time. The data used in this PCA were the ages at which the five molts of each individual took place; the corresponding incremental values are the instar durations, except for the first variable (age at L1–L2 molt) which is identical for both sets (it is also the L1 duration). The top panels show the eigenvalues, expressed as percentages of total variance; note the dominance of the PC1 is even stronger than in the analysis shown in fig 6. The bottom panel shows the coefficients of each instar on the PC1. Standard errors are so small that the symbols entirely cover the error bars.

and the variation in growth curves, and to compare analyses of increments and cumulative trait-at-age data. In addition, the data on instar durations relate this study to the concept of heterochrony. In this discussion, I attempt to synthesize the results of these analyses, linking them to the knowledge on growth processes in hemimetabolous insects and to the evolutionary patterns found in a comparison of ontogenies among all six species of the genus *Limnoporus* (Klingenberg and Spence 1993).

Patterns of Covariation among Morphometric Variables

The analysis with CPCs demonstrates that patterns of static variation are fairly constant throughout postembryonic development. Moreover, these patterns also account for similar proportions of the total variation, and it is mainly the overall amount of variability that increases from early to late instars. As the CPC1 accounts for more than three-quarters of the total variance within each stage, the model of simple allometry fits well in all instars. Similarities of static variation in several instars have also been found in another true bug, the backswimmer *Notonecta maculata* (Cuzin-Roudy 1975), and other water striders (Klingenberg and Zimmermann 1992a). In addition to some constant components of static covariance structure, however, other organisms also show substantial changes associated with key ontogenetic events, for instance those

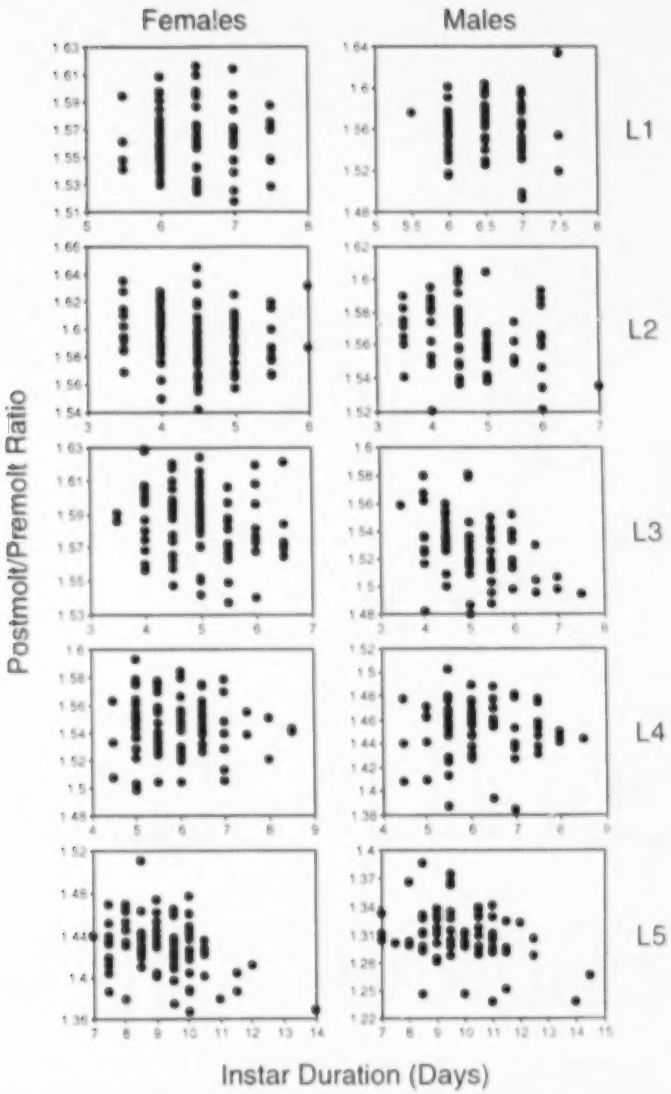


Fig. 6-9. Relations between instar durations and size increments in the following molt. Postmolt/premolt ratios are for the multivariate size variable. The correlations are either very close to zero or negative (especially in the L3 and L5 instars).

that coincide with weaning in rodents (Zelditch 1988; Zelditch and Carmichael 1989).

The patterns of static and ontogenetic allometry are very similar (fig. 6-1). Two explanations for such similarities have been proposed in the literature

(Teissier 1948; Cock 1966; Cheverud 1982b; Klingenberg and Zimmermann 1992a). First, if growth increments along the allometric trajectory, but not in perpendicular directions, are positively correlated with size in the preceding instar, then static and ontogenetic allometry will be more similar in the following instar, and will eventually become identical (Cock 1966). Because the static and ontogenetic patterns tend to be more similar in younger instars (especially the L2) than in later ones, this explanation does not correspond well to the observed patterns. Second, static and ontogenetic allometry are identical if growth vectors between successive stages vary only in their extent, but not their direction (Cheverud 1982b). A less restrictive version of this explanation allows for variation in growth vectors, provided their lengths and directions are uncorrelated (Cheverud 1982b). Regression analyses of the lengths and direction coefficients of growth vectors are only significant for the L3 in females and the L4 in males (after Bonferroni corrections), and R^2 values generally are low. Correlations between direction and length of growth vectors may result from the slight curvature of growth trajectories (Klingenberg and Zimmermann 1992a; Klingenberg and Spence 1993). Although this indicates that there are specific sources of variation during each growth stage, the close overall congruence of static and ontogenetic allometry suggests that variability in the extent of growth contributes most to static variation.

Phenotypic Variation and Constraints in Growth Curves

There are two ways to analyze data measured in each of a number of developmental stages: by stage-specific values and by the increments between successive stages (see also Riska et al. 1984; Lynch 1988; Cowley and Atchley 1992). Although stage-specific values are simply the sum of the value in the first stage and later increments as they accumulate through the growth period, I have shown by direct comparison that analyses of such cumulative and incremental data can produce very different results.

In the analysis of cumulative size, the PC1 strongly dominates (fig. 6-2), suggesting a well-integrated ontogeny and constrained phenotypic variation of growth trajectories. The bulk of the variation affects growth rates in all instars jointly, producing variability in the overall height and in the slopes of growth curves. Only a much smaller fraction (the PC2) is "seesaw" variation contrasting the size in early versus late instars, which gives the growth curves a certain degree of variation in slopes independent of height. The remaining four PCs account for negligible amounts of variation. Conversely, analyses of growth increments suggest there is very little ontogenetic integration, but that increments in different instars are largely independent of each other (figs. 3, 4). In this analysis, not a single PC is an axis of overall growth performance; instead, the first three PCs separately feature variability of growth in the late, middle, and early larval period, respectively (fig. 6-4). Moreover, the PCs differ much less in the amount of variance they account for, and there clearly is variation in all five dimensions. Unlike the cumulative size data, this analysis provides very little evidence for phenotypic constraints in ontogenies. The cumulative and incremental analyses for the CPC2 produced similar differences; this finding therefore applies not only to the "size" variable.

The stark contrast between these two kinds of analyses demonstrates the difference between cumulative and incremental data. In cumulative data, there is a strong ontogenetic autocorrelation, because the value of a variable in a par-

ticular stage is the sum of the value at the preceding stage and the intervening growth increments (see Riska et al. 1984). For instance, if X denotes size in a given instar and Y the growth increment, the covariance between the sizes in this and the following instars is

$$\mathrm{cov}(X, X + Y) = \mathrm{var}(X) + \mathrm{cov}(X, Y).$$

This formula shows that the covariance between successive instars is bound to be high unless there is a large negative covariance between X and Y, which would imply tight regulation of size (see below). In their discussion of ontogenetic autocorrelations, Riska et al. (1984) pointed out that the correlation between successive stages can even become negative if the negative correlations between X and Y are extremely strong or if the variance of Y is much larger than that of X.

In a number of aspects, my findings for cumulative size in water striders remarkably resemble those reported in studies of genetic and phenotypic covariance or correlation matrices for mammals and birds (Cheverud et al. 1983; Leamy and Cheverud 1984; Kirkpatrick and Lofsvold 1989, 1992; Björklund 1993; although the latter three studies used an infinite-dimensional approach, I refer to the eigenfunctions as "PCs" in this comparison). First, the PC1 alone accounts for almost all the total variance, and subsequent PCs for progressively less. Second, the coefficients of the PC1 are positive for all age groups in those studies, indicating joint variation in size at all ages. Furthermore, in some of those examples the PC2s have coefficients that steadily increase or decrease with age in a graded series, with opposite signs at the youngest and oldest ages (see fig. 6-2; Cheverud et al. 1983; Björklund 1993; additional examples have the extreme coefficient values at the second-youngest or second-oldest age, and are thus very close to a graded series, see Kirkpatrick and Lofsvold [1992]). This shows that "seesaw" variation in growth trajectories is not limited to insects.

These similarities for analyses of cumulative data, which range across taxa with such drastically different modes of growth as water striders, rodents, and finches, raise the question whether analyses of incremental data for those other examples would produce as little evidence for ontogenetic integration as for water striders. Published phenotypic and genetic correlation matrices for weight in mice (tables 3, 4 in Riska et al. 1984) provide an opportunity to compare PCA results for cumulative and incremental data directly. For both sexes, and for phenotypic as well as genetic correlations, the PC1 of cumulative data accounts for about three-quarters of the total variance or more, and its coefficients indicate variation in height of the growth trajectory; the PC2s, which take up 9–20% of the total variance, display a "seesaw" pattern of variation. The corresponding analyses of incremental data, however, suggest less severe constraints: the PC1s take up about 25% of the variation in phenotypic, and 50–60% in genetic correlation matrices. The PC coefficients mostly feature contrasts among increments in particular stages; none of them shows easily interpretable patterns like those of the cumulative weights. This congruence between different data sets suggests that the constraints identified in these studies are not a consequence of the underlying physiological or genetic mechanisms, but of the autocorrelation among measurements in successive growth stages.

The fairly tight integration seen in cumulative analyses is straightforward, given the nature of these data. Variation in growth increments at any stage appears again in all subsequent stages, unless subsequent growth compensates for it. Growth variability acting only in the earliest stages therefore contributes to the variation in height of the growth trajectory, in addition to processes affecting growth throughout ontogeny. Likewise, increases or decreases of growth rates mainly at intermediate ontogenetic stages will contribute to "seesaw" patterns. Ontogenetic autocorrelation tends to spread the effects of an ontogenetic event at a particular stage to the successive stages, and thus causes integration and constraints of ontogenetic variation.

Clearly, these constraints on variation are real, even if caused by such ontogenetic autocorrelation, and they can influence the evolution of ontogeny (Cheverud et al. 1983; Kirkpatrick and Lofsvold 1992; Björklund 1993). Because ontogenetic autocorrelation is due to the part-whole relationship inherent in cumulative data and applies to every growth trajectory analyzed in this way, unless there is perfect compensatory growth, the resulting constraints are universal (in the sense of Maynard Smith et al. 1985).

Nevertheless, my study shows that there may be substantial variation in growth increments even in cases where such constraints exist, but this variability may not be apparent in cumulative data as it is swamped with morphometric variation accumulated from earlier stages. The differences in growth trajectories among water striders of the genus *Limnoporus* show that variation was sufficient for size increments in successive instars to evolve relative to each other even in opposite directions (Klingenberg and Spence 1993). This is particularly obvious between *L. canaliculatus* and its sister species *L. esakii* (Andersen and Spence 1992), which grows much more between the L3 and L4, but less between the L4 and L5 instars (fig. 4-5).

Regulation of Growth

A variety of animals have the ability to regulate growth to achieve a particular "target" size, including mammals (e.g., Tanner 1963; Riska et al. 1984; Cameron et al. 1994) and arthropods (e.g., Hartnoll and Dalley 1981; Tanaka 1981; Bryant and Simpson 1984; West and Costlow 1987; Lynch 1988; Freeman 1990). By compensating for variability in earlier stages, whether of genetic or environmental origin, such regulatory growth reduces the variance around the "target" value. The covariance between a measurement in one stage (X) and the subsequent growth increment (Y) plays an important role for determining the variance of the trait in the following stage, because

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + 2\mathrm{cov}(X, Y) + \mathrm{var}(Y).$$

Because var(X) and var(Y) are always positive, a negative cov(X, Y) is the only factor that can keep the morphometric variance constant or even reduce it from one stage to the next (for further discussion, see Riska et al. 1984). The same covariance also relates size regulation to the issue of ontogenetic autocorrelation (see above).

Relative body size, as indicated by the CPC1 scores, is negatively correlated with subsequent growth in the youngest instars (fig. 6-5), indicating convergent growth (Riska et al. 1984). In females, the covariance between the CPC1 scores in the L1 and the following growth increment is just sufficient to main-

tain about the same variance of CPC1 scores in the L2 [i.e., $2\text{cov}(X, Y) \approx -\text{var}(Y)$]. In the L2, this covariance is even negative enough to produce a slight decrease in the variance of CPC1 scores from the L2 to the L3 instar. In males, there is similar decrease in variance between the L1 and L2. The remaining correlations between size and increments up to the L3 instar are also negative, but not significantly different from zero; the corresponding covariances are not sufficiently negative to compensate for the new variance of CPC1 scores produced by variable growth.

In the L4 and L5, correlations between CPC1 scores and increments are positive, indicating divergent growth (Riska et al. 1984), although not significantly different from zero. Despite the relatively low correlations, ranging from 0.04 to 0.21, the corresponding covariances contribute 13–45% of the increase in variance during these instars.

Similar changes occur in several arthropods for which growth regulation has been shown. In the German cockroach, Tanaka (1981) found convergent growth in late, but not in the youngest instars. Several studies for crustaceans have found convergent growth in some stages of development, but not in others (Hartnoll and Dalley 1981; Lynch 1988; Freeman 1990). West and Costlow (1987) found that barnacle larvae tightly regulated growth during the naupliar stages toward target sizes that were unaffected by food concentration, but that there was substantial variation among food treatments in molt increments to the cyprid stage. Likewise, there are drastic changes of growth regulation in mammals, with divergent growth producing large phenotypic and genetic variances in early stages, which decrease later due to convergent growth (Atchley 1984; Riska et al. 1984). Changes in growth regulation therefore are a widespread phenomenon; however, the underlying changes in mechanisms of growth control and the possible adaptive significance are poorly undestood.

Variability in Timing of Molts

The variation in instar durations is substantial, especially in later instars. The slowest-developing bugs can spend up to twice as much time in an instar than the fastest ones (fig. 6-6). Unlike size increments, instar durations are clearly correlated, and show strong integration (fig. 6-7). Therefore, most variation is confined to the overall rate of development, indicating that bugs tend to have either relatively short or relatively long intervals between molts in all instars. This result is especially remarkable given the fact that instar durations are incremental traits; in the corresponding analysis of cumulative ages at the five molts (days after hatching), the phenotypic constraint is even more pervasive due to serial autocorrelation (fig. 6-8).

This large intraspecific variability corresponds to the interspecific variation in the genus *Limnoporus*, which suggests that instar durations have been evolutionarily plastic (Klingenberg and Spence 1993). The species I have studied here, *L. canaliculatus*, is the fastest-developing of the six species; in most instars, the slow-growing *L. genitalis* has instar durations that are about twice as long (Klingenberg and Spence 1993). Even *L. esakii*, the slightly smaller sister species of *L. canaliculatus* (Andersen and Spence 1992), has substantially longer instar durations, especially in the L3. Geographic variation within species (Fairbairn 1984; Firko 1986; Blanckenhorn 1991; Blanckenhorn and

Fairbairn 1995) and interspecific variation (Spence et al. 1980) further underscore the evolutionary flexibility of development time among water striders.

The correlations between instar durations and growth increments were nil to moderately negative (fig. 6-9), and the correlations between total development time and adult size were negative as well, but not significantly different from zero. None of these correlations was positive, as many theoretical models of life history evolution assume, and as they have been found especially in comparisons across larger taxa (Stearns 1992; Roff 1992). While my result contradicts these model assumptions, it is consistent with observations in other water striders species (Blanckenhorn and Fairbairn 1995; Klingenberg and Spence MS) and with the lack of a clear relationship between size and development time among *Limnoporus* species (Klingenberg and Spence 1993).

Heterochronic changes of ontogeny can affect the rates or timing of growth (Alberch et al. 1979; McKinney and McNamara 1991). The possibility to distinguish individual variation in timing of developmental events, such as the molts, from variation in size at a given stage makes hemimetabolous insects excellent systems for studying the evolution of ontogenies. The data from water striders suggest that these parameters vary independently, both within populations (this study) and among species (Klingenberg and Spence 1993). To understand the mechanisms of heterochronic change fully, however, both the genetic covariances and patterns of natural selection would have to be known.

Studies of trait-at-age data have implicitly assumed that animals at a given age are in comparable developmental stages. My results demonstrate that this assumption does not hold generally, but needs to be examined case by case (see also Hall and Miyake 1995). In some instances, the resolution of longitudinal analyses can be improved by adjusting the time axis to a developmental event (e.g., the adolescent growth spurt in humans, Cameron et al. 1994), but even the sequence of developmental events can change among related species (e.g., Strauss 1990), making it difficult to identify homologous stages (see also Alberch 1985). Studies that explicitly consider the timing of multiple developmental events in combination with age-specific size will substantially help to better understand the evolution of ontogeny.

Possible Physiological Mechanisms

My data from water striders, as well as studies in a variety of other organisms, demonstrate size regulation by convergent growth. This regulation occurs by accelerating or slowing growth rates, not by altering instar durations, as there is no positive correlation between instar durations and growth increments (fig. 6-9). The CPC1 and the CPC2 scores both showed targeted growth, but not always in the same instars. Thus regulation seems to affect the overall size of the organism and the relative sizes of its parts in separate ways, suggesting that the corresponding mechanisms act locally.

Intrinsic growth control of individual organs has been shown in the imaginal discs of *Drosophila*, as well as mutations that can disrupt it (Bryant and Simpson 1984). It is not clear, however, how developmental mechanisms such as those of the polar coordinate model (French et al. 1976; Campbell and Tomlinson 1995) contribute to the fine-tuning of larval growth in hemimetabolous insects; experimental tests of these models have focused on qualitative responses

to drastic disruptions of morphological patterns by ablation, grafting, or the misexpression of genes. Two results from physiological studies of insect growth raise further questions on the mechanism for convergent growth. Molting in Heteroptera is induced by stretching of the abdominal wall (Nijhout 1979, 1994), and larger larvae must therefore attain a larger size than smaller ones to trigger a molt. Moreover, the size of the old cuticle and the degree to which it is stretched in the early phase of a molting cycle influence the size in successive instars, because the old cuticle functions as a "template" when the new one is laid down (Bennet-Clark 1971). Both stretch-induced molting and the template effect are more likely to produce divergent growth rather than growth regulation by convergent growth.

The mechanism of stretch-induced molting may partly explain the negative correlations between development time and growth increments (Klingenberg and Spence MS). Bugs with higher growth rates reach the size threshold sooner than slower-growing ones, and therefore initiate the new molting cycle earlier. Because the period from initiation of the molting cycle to the formation of the new cuticle is fairly constant and independent of feeding (Blakley & Goodner 1978), growth that takes place during this period influences the size of the following instar. Consequently, bugs with higher growth rates also increase more in size after initiation of the molting cycle than bugs with slower growth. Together, the earlier molt and larger growth increment of faster-growing individuals lead to a negative correlation between instar durations and growth increments, and to divergent growth.

My data from water striders are only partially consistent with these expected patterns. Clearly, this mechanism does not apply in the youngest instars when there is convergent growth; so it is not surprising that significant negative correlations between instar durations and size increments do not occur before the L3 instar. It is perplexing, however, why the correlations in the L4 instar are so weak, although there are stronger negative correlations in the L3 and L5 instars in both sexes. Moreover, the correlations between total development time and adult size are weak as well.

Overall, these mechanisms of insect growth may account for a part, but clearly not for all the patterns of ontogenetic variation in this data set. This reflects the poor current understanding of the processes controlling postembryonic growth. Although the physiology of growth in Heteroptera is relatively well known, because true bugs have been used as model organisms in this field (mainly *Rhodnius* and *Oncopeltus*), there is only a handful of studies on this topic, and much remains unknown. Clearly, the topics of growth regulation and the correlation between growth in size and development time are promising field for study in both physiology and evolutionary biology.

Conclusions

In this study I simultaneously have considered the covariation of morphometric variables within and among instars, using a model of common principal components specifically suited for longitudinal data (Klingenberg et al. 1996). The CPCs identify components of variation that are independent of each other both within and among instars, and therefore can be analyzed separately without loss of information. This analysis shows that static variation follows con-

stant patterns, and mainly increases in overall magnitude through the growth period.

Analyses of cumulative and incremental data differ dramatically, demonstrating the pervasive influence of serial autocorrelation between successive instars. The patterns of phenotypic variation I found for cumulative growth curves of water striders are similar to those described from studies of both phenotypic and genetic constraints in mammals and birds (Cheverud et al. 1983; Kirkpatrick and Lofsvold 1992; Björklund 1993). These similarities extend further: variation in the overall height of growth trajectories is dominant in all these organisms, and the "seesaw" variation of early versus late stages can be found in most examples. I argue that these are universal constraints due to the part-whole relations inherent in cumulative data; they may severely limit the variation of growth trajectories, but inferences about developmental processes based on such data should be made with caution. In contrast, the analyses of growth increments demonstrate substantial freedom for independent variation in all instars, which corresponds well to the variation observed among *Limnoporus* species (Klingenberg and Spence 1993).

The large variability in instar durations shows that the timing of developmental events differs extensively among individuals. Therefore, age and the instars provide two very different frames of reference for analyzing growth. Individual variability in *L. canaliculatus* corresponds to the interspecific variation found across this genus, where several heterochronic changes have been found (Klingenberg and Spence 1993). Unlike size increments, instar durations are correlated among instars, indicating that individual variation in developmental rates is consistent throughout the larval period.

Overall growth in size switches from a convergent or targeted mode to divergent growth, and as a consequence, the variance in size remains constant during the first three instars, but increases sharply from the L4 to the adult stage. There is a weak negative correlation between instar durations and size increments in some instars, which may be due to the mechanism of stretch-induced molting. This mechanism is particularly interesting: on the one hand, size-triggered molting establishes the basis for independent control of size increments and the timing of molts, but through consistent variation in growth rates, it may produce a correlation between them. Nevertheless, the physiological mechanisms and the genetic basis of ontogenetic variation remain mostly unclear. Much remains to be done to gain an integrated understanding of patterns and processes in the evolution of ontogeny; this study of growth and developmental time is a step toward that goal.

LITERATURE CITED

Airoldi, J.-P., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia). Journal of Zoology (London) 216:21–36.

Alberch, P. 1985. Problems with the interpretation of developmental sequences. Systematic Zoology 34:46–58.

Alberch, P., S. J. Gould, G. F. Oster, and D. B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5:296–317.

Andersen, N. M., and J. R. Spence. 1992. Classification and phylogeny of the Holarctic water strider genus *Limnoporus* Stål (Hemiptera, Gerridae). Canadian Journal of Zoology 70:753–785.

Atchley, W. R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. American Naturalist 123:519–540.

Atchley, W. R., D. E. Cowley, C. Vogl, and T. McLellan. 1992. Evolutionary divergence, shape change, and genetic correlation structure in the rodent mandible. Systematic Biology 41:196–221.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. Biological Reviews of the Cambridge Philosophical Society 66:101–157.

Atchley, W. R., S. Xu, and C. Vogl. 1994. Developmental quantitative genetic models of evolutionary change. Developmental Genetics 15:92–103.

Bennet-Clark, H. C. 1971. The cuticle as a template for growth in *Rhodnius prolixus*. Journal of Insect Physiology 17:2421–2434.

Björklund, M. 1993. Phenotypic variation of growth trajectories in finches. Evolution 47:1506–1514.

Blakley, N., and S. R. Goodner. 1978. Size-dependent timing of metamorphosis in milkweed bugs (*Oncopeltus*) and its life history implications. Biological Bulletin 155:499–510.

Blanckenhorn, W. U. 1991. Life-history differences in adjacent water strider populations: phenotypic plasticity or heritable responses to stream temperature? Evolution 45:1520–1525.

Blanckenhorn, W. U., and D. J. Fairbairn. 1995. Life history adaptation along a latitudinal cline in the water strider *Aquarius remigis* (Heteroptera: Gerridae). Journal of Evolutionary Biology 8:21–41.

Blum, J. W., W. Schnyder, P. L. Kunz, A. K. Blom, H. Bickel, and A. Schürch. 1985. Reduced and compensatory growth: endocrine and metabolic changes during food restriction and refeeding in steers. Journal of Nutrition 115:417–424.

Bryant, P. J., and P. Simpson. 1984. Intrinsic and extrinsic control of growth in developing organs. Quarterly Review of Biology 59:387–415.

Cameron, N., P. Gordon-Larsen, and E. M. Wrchota. 1994. Longitudinal analysis of adolescent growth in height, fatness, and fat patterning in rural South African black children. American Journal of Physical Anthropology 93:307–321.

Campbell, G. and A. Tomlinson. 1995. Initiation of the proximodistal axis in insect legs. Development 121:619–628.

Cheverud, J. M. 1982a. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

———. 1982b. Relationships among ontogenetic, static, and evolutionary allometry. American Journal of Physical Anthropology 59:139–149.

———. 1984. Quantitative genetics and developmental constraints on evolution by selection. Journal of Theoretical Biology 110:155–171.

———. 1995. Morphological integration in the saddle-backed tamarin (*Saguinus fuscicollis*) cranium. American Naturalist 145:63–89.

Cheverud, J. M., J. J. Rutledge, and W. R. Atchley. 1983. Quantitative genetics of development: genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905.

Cock, A. G. 1966. Genetical aspects of metrical growth and form in animals. Quarterly Review of Biology 41:131–190.

Cowley, D. E., and W. R. Atchley. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. American Naturalist 135:242–268.

———. 1992. Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. Evolution 46:495–518.

Creighton, G. K., and R. E. Strauss. 1986. Comparative patterns of growth and development in cricetine rodents and the evolution of ontogeny. Evolution 40:94–106.

Cuzin-Roudy, J. 1975. Étude de la variabilité et de l'allométrie de taille chez *Notonecta maculata* Fabricius (insectes, hétéroptères), par les méthodes classiques et par la méthode des composantes principales. Archives de Zoologie expérimentale et générale 116:173–227.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Ettensohn, C. A., and K. M. Malinda. 1993. Size regulation and morphogenesis: a cellular analysis of skeletogenesis in the sea urchin embryo. Development 119:155–167.

Fairbairn, D. J. 1984. Microgeographic variation in body size and development time in the water strider, *Limnoporus notabilis*. Oecologia 61:126–133.

Firko, M. J. 1986. Comparative life history evolution in the waterstrider, *Gerris remigis*. Ph.D. diss. University of Pennsylvania, Philadelphia.

Flury, B. 1988. Common principal components and related multivariate models. Wiley, New York.

Flury, B. D., and B. E. Neuenschwander. 1995a. Principal component models for patterned covariance matrices, with applications to canonical correlation analysis of several sets of variables. Pp. 90–112 *in* W. J. Krzanowski, ed. Recent advances in descriptive multivariate analysis. Oxford University Press, Oxford.

———. 1995b. Simultaneous diagonalization algorithms with applications in multivariate statistics. Pp. 179–206 in R. Zahar, ed. Approximation and computation: a festschrift in honor of Walter Gautschi. Birkhäuser, Boston.

Freeman, J. A. 1990. Molt increment, molt cycle duration, and tissue growth in *Palaemonetes pugio* Holthuis larvae. Journal of Experimental Marine Biology and Ecology 143:47–61.

French, V., P. J. Bryant, and S. V. Bryant. 1976 Pattern regulation in epimorphic fields. Science 193:969–981.

Good, P. 1994. Permutation tests: a practical guide to resampling methods for testing hypotheses. Springer-Verlag, New York.

Gould, S. J. 1977. Ontogeny and phylogeny. Harvard University Press, Cambridge, Mass.

———. 1989. A developmental constraint in *Cerion*, with comments on the definition and interpretation of constraint in evolution. Evolution 43:516–539.

Hall, B. K. 1992. Evolutionary developmental biology. Chapman and Hall, London.

Hall, B. K., and T. Miyake. 1995. How do embryos measure time? Pp. 3–20 *in* K. J. McNamara, ed. Evolutionary change and heterochrony. Wiley, Chichester.

Hartnoll, R. G., and R. Dalley. 1981. The control of size variation within instars of a crustacean. Journal of Experimental Marine Biology and Ecology 53:235–239.

Jackson, D. A. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology 74:2204–2214.

Kirkpatrick, M., and D. Lofsvold. 1989. The evolution of growth trajectories and other complex quantitative characters. Genome 31:778–783.

———. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.

Klingenberg, C. P. 1996. Multivariate allometry. *In* L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York. In press.

Klingenberg, C. P., B. E. Neuenschwander, and B. D. Flury. 1996. Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components. Systematic Biology. In press.

Klingenberg, C. P., and J. R. Spence. 1993. Heterochrony and allometry: lessons from the water strider genus *Limnoporus*. Evolution 47:1834–1853.

Klingenberg, C. P., and M. Zimmermann. 1992a. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of water striders. American Naturalist 140:601–620.

———. 1992b. Dyar's rule and multivariate allometric growth in nine species of waterstriders (Heteroptera: Gerridae). Journal of Zoology (London) 227:453–464.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.

Leamy, L., and J. M. Cheverud. 1984. Quantitative genetics and the evolution of ontogeny. II. Genetic and environmental correlaions among age-specific characters in randombred house mice. Growth 48:339–353.

Lynch, M. 1988. Path analysis of ontogenetic data. Pp. 29–46 *in* B. Ebenman and L. Persson, eds. Size-structured populations. Springer-Verlag, Berlin.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraints and evolution. Quarterly Review of Biology 60:265–287.

McKinney, M. L., and K. J. McNamara. 1991. Heterochrony: the evolution of ontogeny. Plenum Press, New York.

Monteiro, L. S., and D. S. Falconer. 1966. Compensatory growth and sexual maturity in mice. Animal Production 8:179–192.

Mosimann, J. E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association 65:930–945.

Neuenschwander, B. 1991. Common principal components for dependent random vectors. Ph.D. diss. University of Bern, Switzerland.

Nijhout, H. F. 1979. Stretch-induced moulting in *Oncopeltus fasciatus*. Journal of Insect Physiology 25:277–281.

———. 1994. Insect hormones. Princeton University Press, Princeton, N.J.

Paulsen, S. M., and H. F. Nijhout. 1993. Phenotypic correlation structure among elements of the color pattern in *Precis coenia* (Lepidoptera: Nymphalidae). Evolution 47:593–618.

Pitman, E. J. G. 1937. Significance tests which may be applied to samples from any populations. II. The correlation coefficient test. Journal of the Royal Statistical Society B, 4:225–232.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Riska, B. 1986. Some models for development, growth, and morphometric correlation. Evolution 40:1303–1311.

Riska, B., W. R. Atchley, and J. J. Rutledge. 1984. A genetic analysis of targeted growth in mice. Genetics 107:79–101.

Roff, D. A. 1992. The evolution of life histories: theory and analysis. Chapman and Hall, New York.

Shea, B. T., R. E. Hammer, R. L. Brinster, and M. R. Ravosa. 1990. Relative growth of the skull and postcranium in giant transgenic mice. Genetical Research (Cambridge) 56:21–34.

Spence, J. R., D. H. Spence, and G. G. E. Scudder. 1980. The effects of temperature on growth and development of water strider species (Heteroptera: Gerridae) of central British Columbia and implications for species packing. Canadian Journal of Zoology 58:1813–1820.

Stearns, S. C. 1992. The evolution of life histories. Oxford University Press, Oxford.

Strauss, R. E. 1990. Heterochronic variation in the developmental timing of cranial ossifications in poecilid fishes (Cyprinodontiformes). Evolution 44:1558–1567.

Tanaka, A. 1981. Regulation of body size during larval development in the German cockroach, *Blatella germanica*. Journal of Insect Physiology 27:587–592.

Tanner, J. M. 1963. Regulation of growth in size in mammals. Nature 199:845–850.

Teissier, G. 1948. La relation d'allométrie: sa signification statistique et biologique. Biometrics 4:14–48.

Wayne, R. K., and C. B. Ruff. 1993. Domestication and bone growth. Pp. 105–131 *in* B. K. Hall, ed. Bone. Vol. 7: Bone growth–B. CRC Press, Boca Raton, FL.

West, T. L., and J. D. Costlow. 1987. Size regulation of the crustacean *Balanus eburneus* (Cirripedia: Thoracica). Marine Biology 96:47–58.

Wilson, P. N., and D. F. Osbourn. 1960. Compensatory growth after undernutrition in mammals and birds. Biological Reviews of the Cambridge Philosophical Society 35:324–363.

Zelditch, M. L. 1987. Evaluating models of developmental integration in the laboratory rat using confirmatory factor analysis. Systematic Zoology 36:368–380.

———. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. Evolution 42:28–41.

Zelditch, M. L., and A. C. Carmichael. 1989. Ontogenetic variation in patterns of developmental an functional integration in skulls of *Sigmodon fulviventer*. Evolution 43:814–824.

7. On the Role of Body Size for Life History Evolution*

INTRODUCTION

Body size has been considered traditionally a key determinant of an organism's ecological and physiological properties. Numerous theoretical and empirical studies have explored its connections to other life history traits, such as development time and reproduction (e.g. Peters, 1983; Calder, 1984; Reiss, 1989; Roff, 1992; Stearns, 1992). Most models of life history evolution rest on assumptions about trade-offs between reproductive benefits of size and costs of long growth periods, e.g. through mortality.

A strong positive correlation between development time is commonly assumed, "for one must grow for a longer time to get larger" (Stearns, 1992: 127). Some models for optimal age and size at maturity incorporate size explicitly (Roff, 1981, 1984; Kusano, 1982; Ludwig & Rowe, 1990; Rowe & Ludwig, 1991). Most of these models link development time and adult size either by assuming an allometric relation (Roff, 1981), by using von Bertalanffy growth curves with fixed parameters (Roff, 1984), or by some other growth curve assumed to be constant throughout the population (Kusano, 1982). All these models assume that fecundity is strictly determined by size through an allometric or linear relation (Roff, 1981, 1984; Kusano, 1982; Ludwig & Rowe, 1990; Rowe & Ludwig, 1991). Similar reasoning has been applied in models of latitudinal variation in life-history strategies (Roff, 1980).

Another class of optimization models is based on the assumption that either fecundity increases or offspring mortality decreases as a function of age at maturity (Stearns & Crandall, 1981; Stearns & Koella, 1986; Stearns, 1992). Such fecundity benefits of later maturation are most likely a consequence of size (Fig. 2 in Stearns & Koella, 1986): except in species with parental care where experience results in a reproductive benefit, it is unclear how a delayed first reproduction may lead to an increase in reproductive success unrelated to parental size. Likewise, offspring mortality may decrease as maturity is delayed because larger maternal size increases offspring size and survival (see Stearns & Koella, 1986: 895); again, other explanations that do not involve size assume parental care. Therefore, although Stearns and coworkers do not incorporate size explicitly, their models make crucial implicit assumptions about the role of size, or otherwise only apply to a narrow spectrum of species with parental care. Reaction norms are derived from these models by optimizing age and size at maturity within a series of environments differing in their growth parameters (Stearns & Koella, 1986; Berrigan & Koella, 1994). These models are closely related to those described in the preceding paragraph, despite some biological and mathematical differences, because the optimization is based on a trade-off between fecundity benefits of increased size and associated costs through a prolonged growth period before maturity. Although the models are mostly framed in terms of age, body size plays a central role, and

---

* A version of this chapter is accepted for publication. Klingenberg and Spence 1996. Ecological Entomology.

the assumptions about its relationships to other life history parameters are essential for the applicability of the models.

Despite the crucial importance of these assumptions, surprisingly few studies have critically examined their empirical basis and generality. Studies of vertebrates and laboratory studies of *Drosophila* make up a large part of the data supporting the models (reviewed by Roff, 1992; Stearns, 1992). Other groups have not been investigated comparably and few in-depth studies of invertebrates have been undertaken under field conditions. While a considerable number of studies focus on the correlations between size and either development time or fecundity, most do not assess if the assumptions of life history models are met, as they do not include both aspects simultaneously. Moreover, the published studies may be biased in favour of the models because correlations reported as supporting evidence are often weak, albeit statistically significant, and their biological importance and generality are thus questionable, especially if weak contradictory results are not published.

In this paper, we attempt an empirical validation of the key assumptions made in these models (Oreskes *et al.*, 1994). We report the results from a series of experiments, replicated in several generations, with the water strider *Gerris buenoi* Kirkaldy (Heteroptera: Gerridae). Rearing experiments in the field and several measures of female reproductive performance consistently revealed relationships contradicting those generally assumed in life history models. Our findings show that much of the size variation occurring within natural populations may be selectively neutral. We present an alternative framework for understanding associations among life history traits, in which size does not play an adaptive role.

Material and Methods

Study Organisms

The water strider *Gerris buenoi* Kirkaldy is widespread in North America and occurs abundantly in a range of pond habitats (Spence, 1989). At our study site in central Alberta, Canada, this species has two generations per year. Overwintered adults, which all have fully developed wings, appear on ponds in spring to mate and produce a first ("spring") generation of offspring. The spring generation partly consists of individuals that breed in the same year and do not overwinter, some of which are wingless. These direct breeders produce a second ("summer") generation of offspring, which are invariably long-winged and reproduce only after overwintering.

Study Site and Rearing Experiments

This study was carried out from 1992 to 1994 at the George Lake Field Site, about 100 km northwest of Edmonton in central Alberta, Canada, where we reared *G. buenoi* on two man-made ponds (Experiment Pond in 1992, Meadow Pond in 1993 and 1994; see Spence, 1986, 1989). On these ponds, this species co-occurs naturally with the gerrids *Limnoporus dissortis* (Drake and Harris) and *Gerris comatus* Drake and Hottes.

Most of our results are from individual rearings (only in 1992 a part of the bugs, of the same hatching date, were reared in groups). We placed each larva separately into a bottomless plastic container (diameter ca. 10 cm, rim ca. 6 cm above water), which was kept afloat on the water surface by a ring of plastic foam glued around its outside. Groups of these containers were protected from predators in enclosures screened on all sides (see Spence 1986). In 1993 we experienced problems with discharges of methane gas from the pond sediment, which had accumulated under the bottom screen and occasionally burst up through the enclosure, overturning containers. Thus, 1994 we added a sheet of plastic under each enclosure, separated from the bottom screen by a wooden strut, to ensure that gas bubbles came up outside of the enclosure. Each water strider was checked daily for molts and fed *ad libitum* with frozen insects from a nearby light trap or, if light trap catches were insufficient (summer 1993 only), with frozen flesh flies (*Neobelliera bullata* [Parker]) reared on pork liver in the laboratory.

Within 24 hours after the final molt, we collected the emerging adults, recorded sex and wing morph, and measured the size of each bug. In 1992 and 1993, bugs were killed, dried to constant mass at 60°C, and weighed on a microbalance (resolution: 1 µg). In 1993 and 1994, we measured the total body length to the nearest 0.1 mm.

A separate experiment in the summer generation 1994 examined if our standard rearing procedure caused any biases in the correlations among life history traits, rendering them unrepresentative for the natural population. To assess the effect of the feeding regime, we compared individual rearings with *ad libitum* regime (the "standard" treatment, see above) to a treatment with reduced food levels ("food limitation"), in which each larva received a limited quantity of frozen insects every other day only (e.g. 2–3 chironomid midges; this was severe enough to increase development time and reduce adult size significantly). An additional treatment ("triplet"), with three larvae per container instead of one, allowed social interactions and competition among larvae. Only total body length was measured as a size variable.

To correct for variation in temperature, we calculated development times in degree-days based on air temperatures recorded near the pond surface with a Ryan model 10 recorder (this was not possible in 1992 due to incomplete temperature records, and in the experiment of summer 1994). The underlying linear model of temperature dependence, although it may not be biologically realistic, gives results highly correlated to those from more complex nonlinear models (Lamb, 1992), and is thus appropriate for this study focusing on variation within a population. Because growth thresholds differ between instars (Spence *et al.*, 1980), we computed them separately for each of the five instars (L1–L5) using data from this study population (J. R. Spence, unpubl. data). Thresholds computed by linear regression (± bootstrap standard errors) are 7.4 ± 0.6°C for L1 (n = 58), 10.1 ± 0.5°C for L2 (n = 95), 9.4 ± 0.6°C for L3 (n = 57), 10.0 ± 0.4°C for L4 (n = 71), 9.1 ± 0.3°C for L5 (n = 111). Accumulated degree-days were computed for each of these thresholds from daily minima and maxima, updating values twice daily with a sine-wave interpolation (e.g., Pruess, 1983). Total degree-days are the sum of the values accumulated during each instar.

We computed product-moment (Pearson) correlations between development time and size separately for each generation, sex, and wing morph. We used the bootstrap to compute 95% confidence intervals with the BC_a method (Efron & Tibshirani, 1993). For each confidence interval, we performed 5,000 bootstrap iterations.

Breeding Design and Measures of Female Reproduction

For our experiment in 1992, we caught direct-breeding adults in the field and kept them in a laboratory mass culture for mating and oviposition. Larvae were transferred to field enclosures within two days after hatching. In 1993 and 1994, we used offspring of isolated breeding pairs. For the spring generations, parental bugs were collected in the field as overwintered adults immediately after snowmelt. Mating does not occur in the first few days of activity in spring, ensuring that the breeding females were virgins; none of these females laid fertile eggs when isolated in the laboratory. Larvae were transferred to field enclosures within 24 hours after hatching. We used a randomized list to allocate a position in one of the enclosures to each larva. For the summer generations of 1993 and 1994, we used direct-breeding bugs collected as fifth-instar larvae and immediately separated by sex. Matings were arranged according to a half-sib design (Falconer, 1989). In 1993 we mated each male to two females, and to three females in 1994. Males were switched between the females allocated to them three times weekly (1993) or daily (1994).

Breeding pairs were kept under long-day conditions (19 L: 5 D) in the laboratory at 20°C (1993) or in a climate chamber at the field site at 23°C (1994). Each female or pair was fed a frozen flesh fly three times per week (1993) or either a flesh fly or house fly (*Musca domestica* L.) daily (1994). As these flies are of similar size or larger than the water striders, and remain on the water surface for several days as a potential food source, we consider this an *ad libitum* regime. We provided the females with styrofoam strips for oviposition, which we exchanged at intervals ranging from a week to ten days. For each female, we measured the total body length (to the nearest 0.1 mm) and dry weight after death (only 1993).

We examined the influence of the feeding regime on the relationship between body size and reproduction in an experiment with direct-breeding females in summer 1993. Females were either given a permanent source of food (a frozen flesh fly that was exchanged three times weekly) or a scarce food regime, allowing access to a fly only once a week for eight hours.

Oviposition strips were kept at room temperature in separate containers for hatching. After a period of at least four weeks, which exceeds the incubation time of *G. buenoi* (Spence *et al.*, 1980), we counted the empty eggshells left by hatching larvae and the eggs that failed to develop on all styrofoam strips. Parental bugs were kept in the laboratory until they died; egg counts therefore represent lifetime fecundity. We defined reproductive life span as the time interval from the first egg to the death of the female, because recently laid eggs were present when most females died, and dissections showed that most females still had developing and mature eggs in their ovaries after death. We measured the length and width of 10 eggs from the first two batches of each female (thereby controlling for possible variation of egg size related to female age), and from these, we calculated the volume of the eggs as rotation ellip-

soids, which corresponds fairly well to the shape of gerrid eggs (see also Solbreck *et al.*, 1989).

For correlations involving egg size, we had to adjust the methods to take into account that fewer than 10 eggs were measured for some females. We therefore used weighted correlations, with the number of eggs measured as the weighting factor. Confidence intervals were computed with the bootstrap, using the BC_a method and 5,000 bootstrap iterations (Efron & Tibshirani, 1993).

Estimates of Heritabilities and Genetic Correlations

In the spring generations of 1993 and 1994, we estimated heritabilities of developmental time, total body length, and dry weight (only 1993) of teneral adults. Complete data were available for 513 offspring from 81 dams and 44 sires in 1993, and for 643 offspring from 135 dams and 48 sires in 1994. Data were log-transformed before the analysis to reduce skewness. The statistical model included sex and wing morph as fixed effects, and sire and dam (nested within sire) as random effects. To estimate variance components, we used the restricted maximum likelihood (REML) method (Shaw, 1987), as implemented in the SAS procedure VARCOMP (SAS Institute, 1988). Genetic correlations were computed from the variance components for each character and for their sum. Approximate standard errors for heritabilities and genetic correlations were calculated from the variances and covariances of parameter estimates provided by SAS, using formulas given by Bulmer (1980: 86) and Falconer (1989: 317).

RESULTS

Development Time and Final Size

There was a negative association between development time and adult size in all our experiments (Fig. 7-1). Bugs with a longer larval period tended to be smaller in both body length and dry weight than those that developed more quickly. The estimated correlations between development time and adult size are negative for both size measures, in all experiments, and in each sex and wing morph (Table 7-1). The confidence intervals of the correlation coefficients include zero in only two of the 18 correlations (length, male long-winged bugs in the spring and summer generations of 1993, Table 7-1); all others are significantly negative.

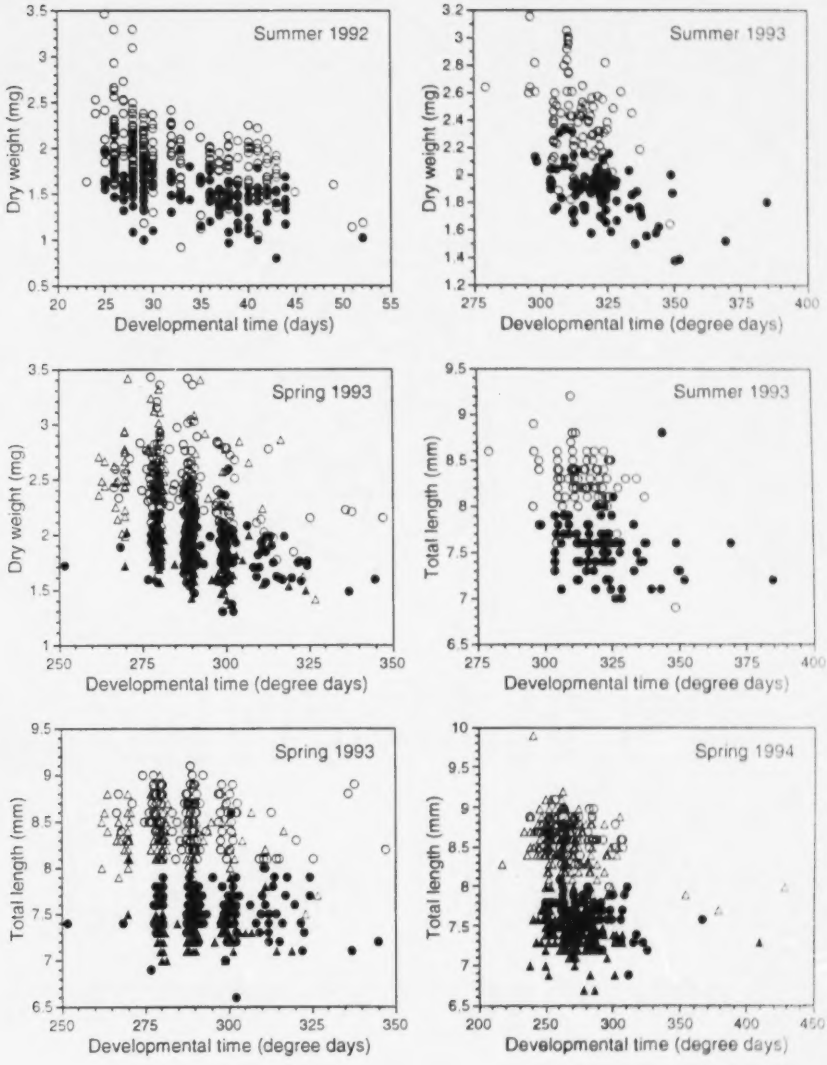


Fig. 7-1. Relation between development time and final size. Size measures are total body length and dry weight of teneral bugs. Solid symbols represent males, open ones females. In the spring generations, there are both long-winged (circles) and wingless bugs (triangles).

Table 7-1. Correlations between development time and adult size. Correlations are Pearson correlation coefficients, and bootstrap 95% confidence intervals were computed with the BC_a method (Efron and Tibshirani 1993). Variables are dry weight (W), total length (L), and development time in days (1992) or in degree-days (1993 and 1994). Abbreviations of wing morphs are lw for long-winged and ap for wingless (apterous).

Size Measure	Sex	Wing Morph	Correlation Coefficient	Confidence Interval	Sample Size
Summer generation 1992:					
W	f	lw	–.44	[–.53, –.35]	217
W	m	lw	–.50	[–.61, –.37]	149
Spring generation 1993:					
W	f	lw	–.37	[–.47, –.24]	145
W	f	ap	–.42	[–.59, –.18]	125
W	m	lw	–.26	[–.38, –.12]	168
W	m	ap	–.42	[–.59, –.19]	75
L	f	lw	–.23	[–.40, –.02]	145
L	f	ap	–.36	[–.56, –.08]	125
L	m	lw	–.05	[–.22, .14]	168
L	m	ap	–.33	[–.53, –.09]	75
Summer Generation 1993:					
W	f	lw	–.41	[–.58, –.19]	83
W	m	lw	–.50	[–.64, –.32]	88
L	f	lw	–.37	[–.62, –.09]	83
L	m	lw	–.17	[–.39, .10]	87
Spring generation 1994:					
L	f	lw	–.30	[–.49, –.06]	54
L	f	ap	–.37	[–.49, –.22]	249
L	m	lw	–.29	[–.49, –.08]	101
L	m	ap	–.19	[–.28, –.09]	239

Heritabilities (± standard errors) computed from half-sib correlations, i.e. using the sire component to estimate additive variances, were 0.50 ± 0.20 for length, 0.11 ± 0.13 for dry weight, and 0.26 ± 0.14 for development time in 1993. The corresponding genetic correlations were 0.11 ± 0.55 between dry weight and development time and -0.18 ± 0.32 between body length and development time. Genetic correlations computed from full-sib correlations, which also include dominance and maternal effects, were -0.43 ± 0.25 and -0.39 ± 0.18, and thus corresponded more closely to the phenotypic correlations. These estimates, however, have to be interpreted cautiously, as suggested by the large standard errors, and because the REML analysis produced a zero estimate for the between-dam component of the sum of length and development time. Despite the larger sample size and increased number of families in 1994, a zero value was estimated for the between-sire component of body length, and half-sib correlations therefore could not be used for this variable.

The heritability of development time based on half-sib correlation was 0.18 ± 0.13, and the heritabilities based on full-sib correlations were 0.31 ± 0.08 for length and 0.29 ± 0.07 for development time. The genetic correlation, also based on full-sib correlations, was –0.46 ± 0.14, and therefore similar to the estimate of the previous year. A combined analysis of both years, with year as an additional fixed effect, gave results very similar to those from the 1994 experiment alone.

Phenotypic correlations between development time and adult size in water striders reared under food limitation or in small groups were similar to those in bugs from the standard rearing procedure (Table 7-2). The negative correlations between development time and adult size are therefore not an artifact of the specific experimental conditions, but can be expected to apply in free-living water striders as well.

Table 7-2. Effect of food limitation and competition on correlations between development time and adult size. Development time is measured in days, and the size variable is total body length. Water striders in the standard treatment were reared in isolation and fed daily; those under food limitation were only fed every other day; and in the "triplet" treatment, three bugs were reared in each container, thus allowing social interactions and competition among them. Correlations are Pearson correlation coefficients, and bootstrap 95% confidence intervals were computed with the BCa method (Efron and Tibshirani 1993).

Sex	Correlation Coefficient	Confidence interval	Sample size
Standard:			
f	–0.79	[–0.92, –0.35]	30
m	–0.60	[–0.84, –0.12]	43
Food limitation:			
f	–0.53	[–0.69, –0.27]	21
m	–0.55	[–0.74, –0.15]	23
Triplet:			
f	–0.59	[–0.76, –0.35]	87
m	–0.46	[–0.62, –0.24]	80

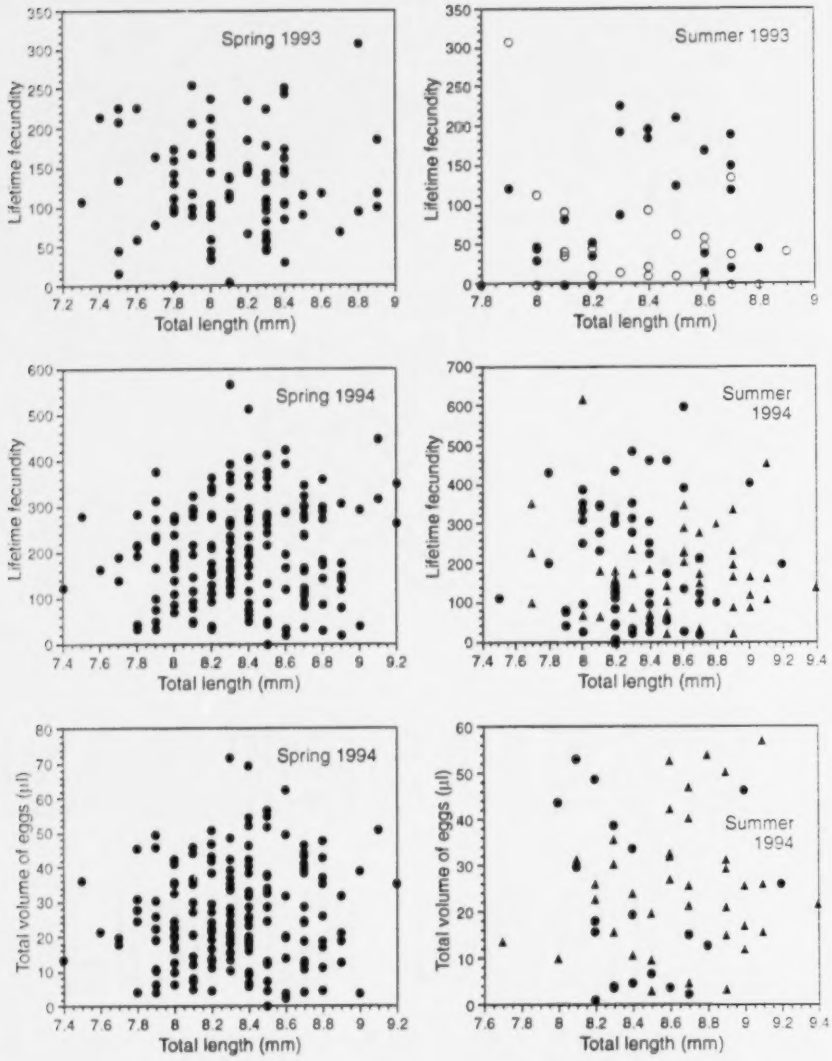


Fig. 7-2. Relation between female size and measures of reproductive output. Reproductive output is characterized by lifetime fecundity (top four panels) and the estimated total volume of eggs laid by a female (only in 1994; bottom row). In the experiment of summer 1993, some females were fed *ad libitum* (solid circles) and others at a reduced level (open circles). For the 1994 summer generation, circles stand for long-winged females, and triangles for wingless females.

Female Size and Reproduction

We studied several measures of female reproductive performance and their relations to the size of the female. Lifetime fecundity shows no relation to female size (Fig. 7-2). The differences in lifetime fecundity between females of the same generation in different years may reflect differences in laboratory conditions. Correlation coefficients are close to zero in all generations and wing morphs, and confidence intervals are fairly narrow and include zero (Table 7-3). As another measure of total reproductive effort, we estimated the total volume of the eggs laid by a female as her lifetime fecundity multiplied by the average volume of her eggs. In the experiment of spring 1994, the scatter of total egg volume versus female length closely resembles the pattern for lifetime fecundity, but less so in summer (Fig. 7-2). Again, correlations are low and all confidence intervals include zero (Table 7-3). These correlations indicate that these two measures of reproductive effort are independent of a female's body size.

Table 7-3. Correlations between female size and reproduction. Correlations are Pearson correlation coefficients, and bootstrap 95% confidence intervals were computed with the BC_a method (Efron and Tibshirani 1993). Size measures are total length (L) and dry weight (W) after natural death of the female. Abbreviations of generations are OW for overwintered (diapause) and DB for direct-breeding bugs, and wing morphs are denoted lw for long-winged and ap for wingless (apterous).

Size Measure	Generation	Wing Morph	Year	Correlation Coefficient	Confidence Interval	Sample Size
Lifetime fecundity:						
L	OW	lw	1993	.002	[–.24, .18]	89
W	OW	lw	1993	.15	[–.06, .35]	84
L	OW	lw	1994	.09	[–.06, .22]	179
L	DB	lw	1994	.05	[–.19, .27]	61
L	DB	ap	1994	–.03	[–.32, .33]	46
Reproductive life span:						
L	OW	lw	1993	–.06	[–.27, .18]	81
W	OW	lw	1993	–.14	[–.33, .05]	80
L	OW	lw	1994	–.06	[–.19, .08]	176
L	DB	lw	1994	–.27	[–.43, –.08]	60
L	DB	ap	1994	–.44	[–.64, –.14]	46
Average volume per egg:*						
L	OW	lw	1994	–.01	[–.17, .15]	163
L	DB	lw	1994	–.11	[–.44, .25]	20
L	DB	ap	1994	–.07	[–.36, .20]	38
Total volume of eggs:*						
L	OW	lw	1994	.13	[–.02, .26]	163
L	DB	lw	1994	–.14	[–.60, .39]	20
L	DB	ap	1994	.17	[–.09, .41]	38
Proportion of eggs hatched:†						
L	OW	lw	1993	.16	[–.09, .36]	87
W	OW	lw	1993	.28	[.05, .43]	82
L	OW	lw	1994	–.03	[–.21, .13]	164
L	DB	lw	1994	.35	[.03, .58]	36
L	DB	ap	1994	.09	[–.23, .36]	41

* For traits involving egg measurements, correlations are weighted by the number of eggs measured for each female.

† Breeding pairs without any hatched eggs are excluded to ensure all females were mated to a fertile male.

Likewise, reproductive life span appears to be independent of female size. In summer 1994, however, large females tended to survive for only a relatively short time (Fig. 7-3), giving rise to negative correlations for both wing morphs (Table 7-3). These contrast with the very weak correlations found in the other two generations (with larger sample sizes), and therefore should be interpreted cautiously.

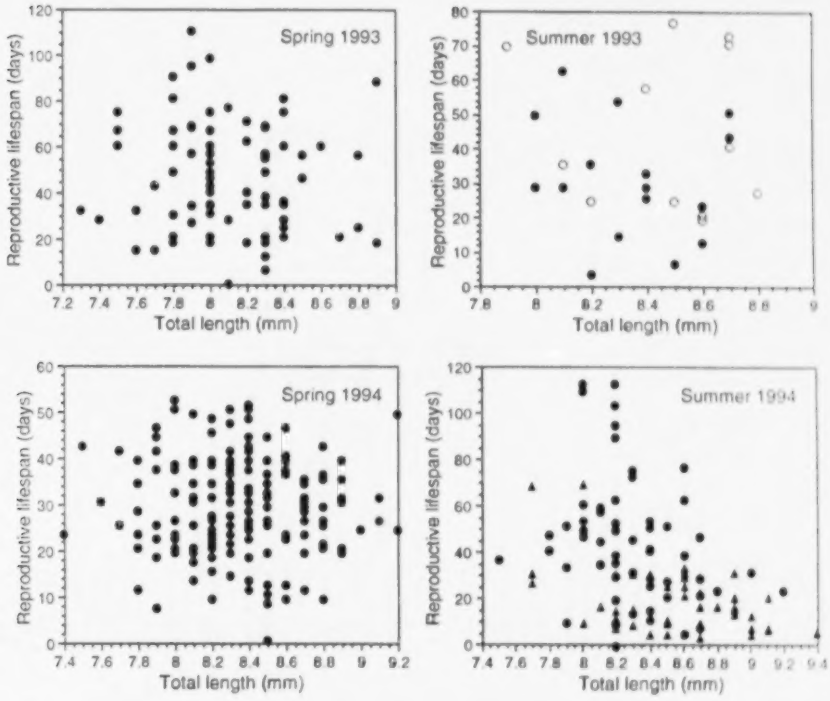


Fig. 7-3. Relation between female size and reproductive life span. Symbols are as in Fig. 7-2.

To assess if larger size may benefit females through the quality of their eggs, we considered egg volume and the proportion of eggs hatched (Fig. 7-4). The volume per egg appears to be unrelated to female size, as correlation coefficients are close to zero and all confidence intervals include zero (Table 7-3). There are two cases where the correlations of the proportion of eggs hatched with size are statistically significant. These are the spring experiment 1993, with dry weight as a size variable, and the long-winged bugs in summer 1994. Conversely, the largest sample (spring 1994) produced a correlation very close to zero. Although these data suggest that larger size is correlated with the proportion of fertile eggs in some replicates, this correlation does not hold generally for any size measure, generation or wing morph (also note that there are 21 correlations in Table 7-3, and that we did not apply a Bonferroni adjustment of significance levels).

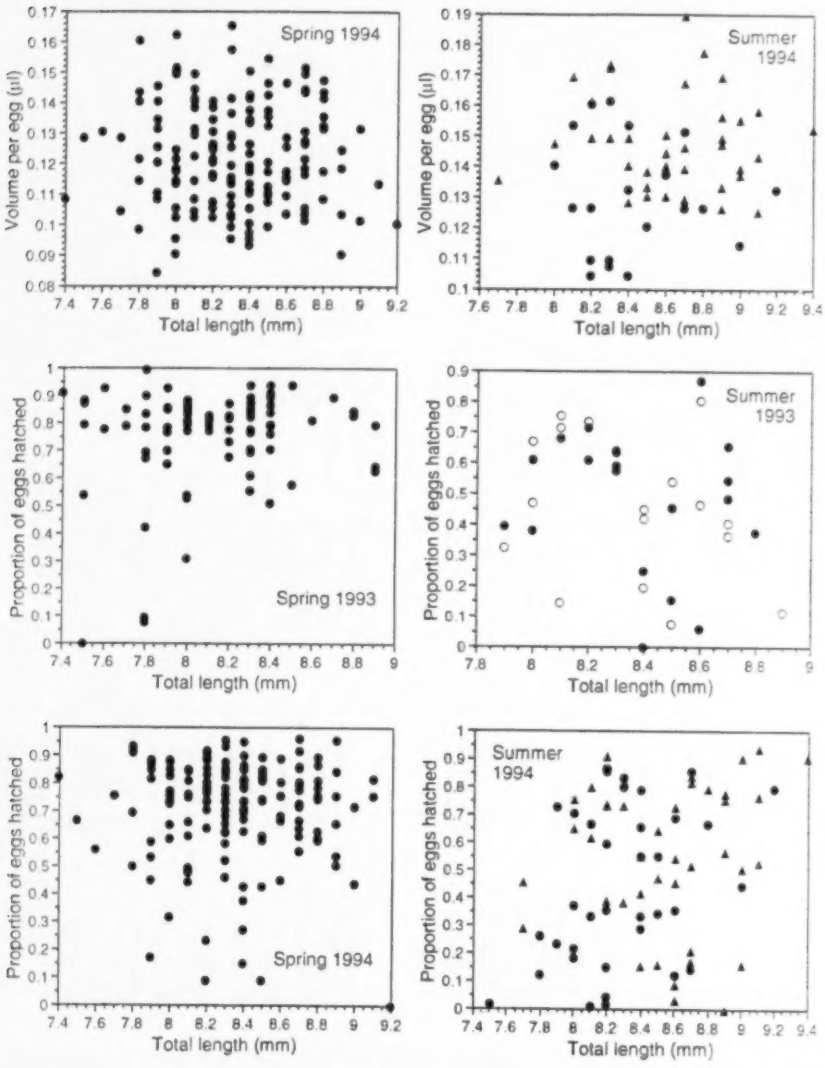


Fig. 7-4. Relation between female size and measures of egg size and quality. Egg size is the average volume in a sample of eggs from each female (top two panels). Egg quality is given as the proportion of each female's eggs that hatched (bottom four panels). Symbols are as in Fig. 7-2.

We also considered the possibility that the experiments described above missed an important benefit of body size, as larger size of females may improve their reproductive performance only under food stress, but not under an *ad libitum* feeding regime. To examine the importance of nutritional status on female reproductive traits, we carried out a small experiment using direct

Table 7-4. Influence of food levels on correlations between female size and reproduction. At the high food level, females had permanent access to food (a frozen flesh fly renewed three times per week), whereas at the low food level, they had the opportunity to feed only for eight hours per week. Correlations are Pearson correlation coefficients, and bootstrap 95% confidence intervals were computed with the BC_a method (Efron and Tibshirani 1993). Size measures are total length (L) and dry weight (W) after natural death of the female.

	High Food Level (*ad libitum*)			Low Food Level		
Size Measure	Correlation Coefficient	Confidence Interval	Sample Size	Correlation Coefficient	Confidence Interval	Sample Size
Lifetime fecundity:						
L	.12	[−.26, .52]	21	−.35	[−.65, .30]	22
W	.52	[.14, .77]	21	.10	[−.72, .58]	22
Reproductive life span:						
L	−.11	[−.62, .40]	13	−.15	[−.63, .40]	16
W	.33	[−.28, .78]	13	−.66	[−.90, −.20]	16
Proportion of eggs hatched:*						
L	−.13	[−.49, .22]	19	−.31	[−.81, .18]	18
W	−.24	[−.53, .09]	19	−.08	[−.46, .32]	18

* Breeding pairs without any hatched eggs are excluded to ensure all females were mated to a fertile male.

breeding females in summer 1993. Under food stress, correlations of size with any of the reproductive traits were not or only slightly higher (more positive) than in the high food regime (Table 7-4). Contrary to our expectations, several of the correlations even were lower (more negative) when food was scarce. While the small sample sizes make it difficult to assess the biological importance of the latter result, the experiment clearly showed that there are no benefits of larger body size under a low food regime.

Discussion

Our experiments, replicated in several generations in the same population, showed a consistent negative correlation between development time and adult size but did not reveal any clear relationship between female size and several measures of reproductive performance. The fairly narrow confidence intervals clearly indicate that our samples were sufficiently large to provide adequate statistical power for characterizing phenotypic correlations.

The negative correlations between development time and adult size suggest that there is unexpected variation in overall vigour. Some individuals have a substantially shorter larval period and also tend to grow larger than others. The additive genetic correlations seem to be in line with the phenotypic ones, although the large standard errors suggest that these estimates are unreliable.

Genetic correlations estimated from full-sib correlations, which include some of the variation due to dominance and maternal effects, closely match the phenotypic correlations. The absence of strong genetic correlations provides flexibility for independent evolutionary changes of adult size and development time, as it has been observed in comparisons among several water strider species (Fairbairn, 1990; Klingenberg and Spence, 1993).

Negative correlations between development time and adult size have often been reported for experiments in which animals were exposed to a range of environments varying in suitability, e.g., different food regimes, both in water striders (Blanckenhorn, 1994) and other insects (e.g., Dixon, 1985; Gebhardt & Stearns, 1988, 1993; Solbreck *et al.*, 1989; Roff, 1992; Stearns, 1992; Panizzi & Saraiva, 1993; Sota, 1993). Variation in food availability, however, cannot the source of negative correlations in our experiments, because correlations were calculated only within experimental treatments; for each correlation estimate, all larvae were either fed *ad libitum* or the same reduced regime. As the larvae for each experimental replicate were reared simultaneously on the same pond, other environmental factors are unlikely to have a major effect on variation within generations. Moreover, these negative correlations are not a result of the favourable conditions used in our standard rearing procedure, but they also occur in experiments under food limitation and in group-reared larvae. Because this range of experimental conditions approximates those found in natural habitats of this species, we expect negative correlations between development time and final size in natural populations. Therefore, our results indicate that the water striders' intrinsic variation in growth rates far exceeds the influence of a possible trade-off (van Noordwijk & de Jong, 1986; Houle, 1991).

In experiments with another species of water striders, Blanckenhorn & Fairbairn (1995) found a similar negative correlation between development time and adult size within populations. Within constant environments, negative correlations between development time and final size also have been found in *Drosophila* (Gebhardt & Stearns, 1993), in the southern green stink bug (McLain, 1991), and most clearly in breeding and selection experiments with milkweed bugs (Hegmann & Dingle, 1982; Palmer & Dingle, 1986; Dingle *et al.*, 1988). These studies show that variation in overall vigour, rather than a trade-off between development time and adult size, is also dominant in other species.

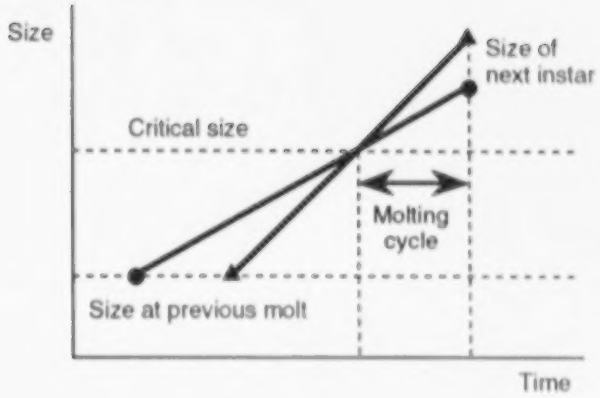


Fig. 7-5. Size-triggered molting as an explanation for the negative correlation between development time and final size. The bold lines represent growth of two individuals, shifted along the time axis so that they attain the critical size simultaneously, and along the size axis so they have the same initial size. Logarithmic transformation of the size axis accommodates multiplicative growth and the mechanism of molting control, which responds to abdominal stretch or internal pressure. The negative correlation results because the slower-growing individual (*dots*) takes longer to reach the critical size than the faster-growing one (*triangles*), which also has the larger size increment because it grows at its higher rate during the molting cycle (i.e. from reaching critical size to ecdysis). Because the fast-growing insect enters the next instar with a larger size, the gaps in development time and size widen in every instar, as long as the growth rates, the duration of a molting cycle, and the properties of stretch or pressure receptors remain constant.

The physiological processes of molting control in insects may cause or at least reinforce these negative correlations. In Heteroptera, which are particularly well studied, and in other insects, the initiation of molting cycles is strongly size-dependent and occurs only after a larva has reached a critical size (e.g., Blakley & Goodner, 1978; Nijhout, 1979, 1994; Woodring, 1983; Sehnal, 1985). Moreover, this size threshold itself depends on the size at the outset of the instar, because the new molting cycle is triggered by the abdominal distention resulting from growth of internal organs during the instar (Nijhout, 1979). Because this mechanism most likely involves receptors responding to stretch (Nijhout, 1984) or pressure on the body wall (Chiang & Davey, 1988), the critical size will be a multiple of the size at the start of the instar (examples in Nijhout, 1994). Conversely, the temporal sequence and duration of the molting cycle (from reaching critical size to ecdysis) is fairly constant irrespective of growth rate or feeding regime (Bennet-Clark, 1971; Blakley & Goodner, 1978; Woodring, 1983). As the period after the initiation of the molting cycle can take up a substantial portion of the whole instar duration, growth during this time is a major determinant of size in the following instar (Bennet-Clark, 1971; Blakley & Goodner, 1978; Nijhout, 1979; Woodring, 1983). More importantly, the old cuticle serves as a template when the new epicuticle is laid down, and stretch by food or the growing internal organs therefore leads to increased size of the next instar (Bennet-Clark, 1971). Together, stretch-induced initiation and constant duration of the molting cycle

can generate a negative correlation between development time and size. Individuals with a slower growth rate attain the critical size later and thus have a longer instar duration, but those that grow faster have a larger size increment during the instar because they grow more between reaching the size threshold and the actual molt (Fig. 7-5). If this process occurs in all instars, the negative correlations it generates between development time and size are expected to be stronger from instar to instar. Whereas it is likely that this explanation applies to the correlations reported for species of *Oncopeltus* (Hegmann & Dingle, 1982; Palmer & Dingle, 1986; Dingle *et al.*, 1988) for which physiological studies were done, it is unclear if it can be extrapolated to water striders, because the only other Heteroptera studied are two species of reduviids (Nijhout, 1984, 1994). The physiology of molting control, however, is not the only possible explanation for our observations. Blakley (1981) searched for adaptive explanations of size-triggered molting itself, and the studies of Ludwig & Rowe (1990) and Rowe & Ludwig (1991) explained a negative correlation between development time and final size as the optimal reaction norm resulting from a trade-off between emergence time versus final size and fecundity. Instead, we suggest this physiological mechanism as a possible alternative to explain the observed life history patterns through intrinsic properties of the organisms, viewing it as a developmental "constraint" rather than invoking adaptation (Gould & Lewontin, 1979; Maynard Smith *et al.*, 1985).

All five measures of reproductive performance that we considered appear to be unrelated to female size. The highest correlations tended to occur in replicates with relatively small samples (Table 7-3). Given that we computed 21 correlations, the four for which the 95% confidence intervals did not include zero should be interpreted cautiously. Besides this concern about statistical significance, there is the question of biological importance of these associations. Even in the replicate with the highest correlation coefficient (between female length and the proportion of eggs hatched for long-winged direct breeders in 1994, Table 7-3), size would only account for approximately 12% of the total variability in that reproductive trait. Any trade-off based on these relationships would be equally weak.

These findings closely agree with other data from *Gerris buenoi*. Rowe & Scudder (1990) found that body size accounted for less than 5% of the variation of several reproductive traits including lifetime fecundity. Fairbairn (1988) reported positive correlations between female body length and the number of eggs carried in the ovaries for *G. buenoi* and two other water striders. Although these were statistically significant, body size explained 11% of the variation or less, and it is therefore unclear how important body size is relative to other factors. In *Aquarius remigis*, Firko (1986) found some strong correlations of body size with lifetime fecundity and reproductive life span, but these were so inconsistent between food treatments or study populations that he concluded size did not account for a significant portion of the variation (p. 95); Blanckenhorn (1994), Blanckenhorn & Fairbairn (1995), and Blanckenhorn *et al.* (1995) reported similar results.

Because a female's greater body reserves may be the main advantage of larger size, experiments under *ad libitum* conditions may not reveal size effects important under food limitation. Our experiment with two different food levels indicates that size does not provide a reproductive benefit even under food

stress (Table 7-4). This agrees with the results of similar experiments reported by Firko (1986), Rowe & Scudder (1990), and Blanckenhorn (1994). Because the method of food delivery in the low food treatment differed from study to study (e.g., small amounts of food permanently available vs. large amounts accessible for a restricted time), the temporal pattern of food availability can be ruled out as a factor itself. The lack of a clear positive correlation between size and reproductive performance is therefore real, and not an artifact of the *ad libitum* regime used in our experiments.

Because we only estimated phenotypic, but not genetic correlations between size and reproductive performance, one might suppose that a positive genetic correlation could exist. This implies that the positive genetic covariances are compensated by similarly large negative environmental covariances, to add up to our observed phenotypic covariances near zero. In more concrete terms, this means that individuals with an environmentally induced size advantage (e.g. by better nutrition) would suffer a penalty of reduced reproduction. We cannot imagine any physiological or ecological mechanism that might produce such a negative environmental covariance, and therefore rule out this possibility.

A number of studies in other arthropods are consistent with our results and report the lack of correlations between female size and fitness components, such as lifetime fecundity (Slansky, 1980; Boggs, 1986; Leather & Burnand, 1987; Johnson, 1990; Spence *et al.*, 1996), female life span (Slansky, 1980; Boggs, 1986; Leather & Burnand, 1987), and egg size (Boggs, 1986 [except for unusually large or small females]; Leather & Burnand, 1987; Solbreck *et al.*, 1989). Larsson (1989) even found negative correlations of female size with fecundity and survival. In the majority of published studies, however, correlations between female size and female fitness components are positive, including those for fecundity (e.g., Solbreck *et al.*, 1989; Marshall, 1990; Kasule, 1991; Honek, 1993; Messina, 1993), longevity (e.g., McLain, 1991), and egg size (e.g., Marshall, 1990; Yafuso, 1994). It is difficult to assess the generality and biological importance of these relationships, as the correlations are variable, although most are stronger than those reported for water striders (see above). We suspect that the literature, largely from applied entomology, is biased to a certain degree against cases where reproduction is independent of size: if there is no need to consider size as a covariate, authors may not even mention it at all. Moreover, various reproductive traits are related to each other in ways that can be complex and sometimes counterintuitive; for instance, the proportion of eggs hatched declines with increasing egg size in the moth *Parapediasia teterrella* (Marshall, 1990). Altogether, the relationships between size and reproductive traits reflect the ecological and physiological diversity of insects; robust generalizations are unlikely (see also Leather, 1988).

A referee of an earlier version of this paper suggested that sexual selection, especially in males, might constitute a benefit for individuals of larger size. Yet, the literature on sexual selection in waterstriders is ambiguous or openly contradictory (reviewed by Arnqvist, 1996). In *Gerris buenoi*, Fairbairn (1988) found significant mating advantages of both large *and* small males in replicate samples, but no overall effect; likewise, there was no significant overall mating advantage for large females. In *Aquarius remigis*, the best-studied species, some authors reported no significant overall effect of size (Rubenstein, 1984; Fairbairn, 1988), another study found sexual selection favouring large males

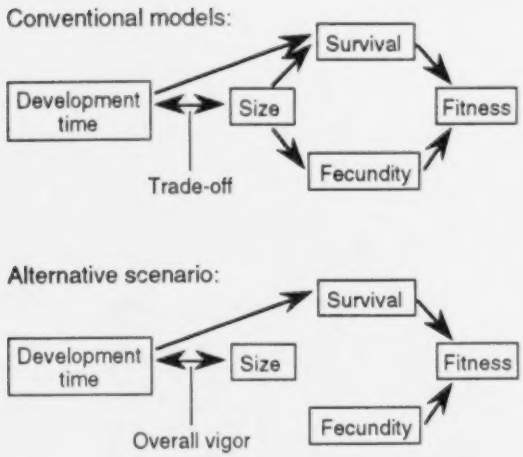


Fig. 7-6. Relationships among life history traits in conventional models and our alternative scenario for *Gerris buenoi*. Size is central to conventional optimization models, because they depend on a trade-off between development time and size that is coupled to reproductive benefits of increased size. Moreover, survival may also be a function of size, although this is often not incorporated in life history models. In *G. buenoi*, size is unrelated to reproductive traits, and is only correlated to development time through variation in overall vigour of larvae.

(Fairbairn & Preziosi, 1994), whereas Fairbairn (1993) and Blanckenhorn *et al.* (1995) reported selective advantages of small males. Thus, the published evidence does not suggest any clear advantage through sexual selection for large gerrids.

Our results indicate that the relationships of size with other life history traits in *Gerris buenoi* contradict those assumed in conventional models and therefore suggest an alternative scenario for life history evolution (Fig. 7-6). Whereas the optimality models for age and size at maturity assume a trade-off between development time and adult size, our data suggest that the correlation between them is due to variation in overall vigour and possibly coupled with the effects of stretch-induced molting. In *G. buenoi*, we could not find the positive relationship between size and several reproductive traits that provides the benefits of larger size in optimality models. Within the limits of natural variation in our study population, size therefore appears to be selectively neutral, rather than an adaptive trait. It may evolve as a correlated response to selection for dispersal (Vepsäläinen, 1978; Spence, 1989), with which it may be positively correlated (Dingle, 1991), or for short development time, although low genetic correlations should limit this latter possibility in our system. In contrast, the variation among species reflects adaptive evolution, as size and life histories of water striders are correlated with habitat use (e.g., Vepsäläinen, 1978; Spence, 1981, 1983, 1989; Andersen, 1982: 331–344).

We used the relations among life history traits for a natural population of *G. buenoi*, combined with information on physiological processes in other Het-

eroptera, to suggest an alternative scenario for life history evolution. Clearly, such an alternative will not supplant the models now accepted for well-studied organisms such as *Drosophila*, fish, and mammals. We expect, however, that numerous patterns of associations and trade-offs among life history traits occur in nature, and that additional variables may play a role besides those commonly studied. Disease and parasites may be such factors; in our study population, however, preliminary analyses showed that they cannot account for the discrepancies between our results and conventional life history models (unpubl. data). To explore this diversity, detailed empirical studies of additional species are needed that include multiple life history traits simultaneously.

Moreover, our results suggest that the range of theoretical models should be expanded to include new combinations of correlations and constraints, as well as neutral evolution. This opens the theory of life histories, currently an exclusive domain of adaptationist models (Partridge & Harvey, 1988; Roff, 1992; Stearns, 1992), to the debate about adaptation and optimality (Gould & Lewontin, 1979; Reeve & Sherman, 1993; Orzack & Sober, 1994). Optimization models consider an adaptive landscape with "peaks" corresponding to optimal phenotypes, "ridges" depicting optimal reaction norms, or gentle ascents towards a boundary set by the assumption of a trade-off. Instead, our data suggest the adaptive landscape has a flat "plateau." On this "plateau" all phenotypes have approximately equal fitnesses, whereas natural selection presumably occurs on the surrounding "slopes," which are beyond the bounds of variation in our study population. Delimiting such domains of neutral and of adaptive evolution and characterizing their spatial and temporal variation will broaden the scope of both empirical and theoretical life history studies.

REFERENCES

Andersen, N.M. (1982) The Semiaquatic Bugs (Hemiptera, Gerromorpha): Phylogeny, Adaptations, Biogeography, and Classification. Scandinavian Science Press, Klampenborg, Denmark.

Arnqvist, G. (1996) The evolution of water strider mating systems: causes and consequences of sexual conflict. In: Social Competition and Cooperation in Insects and Arachnids. Vol. I. Evolution of Mating Systems (ed. by J. C. Choe & B. J. Crespi). Princeton University Press. In press.

Bennet-Clark, H.C. (1971) The cuticle as a template for growth in *Rhodnius prolixus*. Journal of Insect Physiology, **17**, 2421–2434.

Berrigan, D. & Koella, J.C. (1994) The evolution of reaction norms: simple models for age and size at maturity. Journal of Evolutionary Biology, **7**, 549–566.

Blakley, N. (1981) Life history significance of size-triggered metamorphosis in milkweed bugs (*Oncopeltus*). Ecology, **62**, 57–64.

Blakley, N. & Goodner, S.R. (1978) Size-dependent timing of metamorphosis in milkweed bugs (*Oncopeltus*) and its life history implications. Biological Bulletin (Woods Hole), **155**, 499–510.

Blanckenhorn, W. U. (1994) Fitness consequences of alternative life histories in water striders, *Aquarius remigis* (Heteroptera: Gerridae). Oecologia (Berlin), **97**, 354–365.

Blanckenhorn, W.U. & Fairbairn, D.J. (1995) Life history adaptation along a latitudinal cline in the water strider *Aquarius remigis* (Heteroptera: Gerridae). Journal of Evolutionary Biology, **8**, 21–41.

Blanckenhorn, W.U., Preziosi, R.F. & Fairbairn, D.J. (1995) Time and energy constraints and the evolution of sexual size dimorphism—to eat or to mate? Evolutionary Ecology, **9**, 369–381.

Boggs, C.L. (1986) Reproductive strategies of female butterflies: variation and constraints on fecundity. Ecological Entomology, **11**, 7–15.

Bulmer, M.G. (1980) The Mathematical Theory of Quantitative Genetics. Clarendon Press, Oxford.

Calder, W.A., III. (1984) Size, Function, and Life History. Harvard University Press, Cambridge, Mass.

Chiang, R.G. & Davey, K.G. (1988) A novel receptor capable of monitoring applied pressure in the abdomen of an insect. Science, **241**, 1665–1667.

Dingle, H. (1991) Evolutionary genetics of animal migration. American Zoologist, **31**, 253–264.

Dingle, H., Evans, K.E. & Palmer, J.O. (1988) Responses to selection among life-history traits in a nonmigratory population of milkweed bugs (*Oncopeltus fasciatus*). Evolution, **42**, 79–92.

Dixon, A.F.G. (1985) Aphid Ecology. Blackie, Glasgow.

Efron, B. & Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Chapman and Hall, New York.

Fairbairn, D.J. (1988) Sexual selection for homogamy in the Gerridae: an extension of Ridley's comparative approach. Evolution, **42**, 1212–1222.

Fairbairn, D.J. (1990) Factors influencing sexual size dimorphism in temperate waterstriders. American Naturalist, **136**, 61–86.

Fairbairn, D.J. (1993) The costs of loading associated with mate-carrying in the waterstrider, *Aquarius remigis*. Behavioral Ecology, **4**, 224–231

Fairbairn, D.J. & Preziosi, R.F. (1994) Sexual selection and the evolution of allometry for sexual size dimorphism in the water strider, *Aquarius remigis*. American Naturalist, **144**, 101–118.

Falconer, D.S. (1989) Introduction to Quantitative Genetics. 3rd ed. Longman, New York.

Firko, M.J. (1986) Comparative life history evolution in the water strider, Gerris remigis. Ph.D. diss. University of Pennsylvania.

Gebhardt, M.D. & Stearns, S.C. (1988) Reaction norms for developmental time and weight at eclosion in *Drosophila mercatorum*. Journal of Evolutionary Biology, **1**, 335–354.

Gebhardt, M.D. & Stearns, S.C. (1993) Phenotypic plasticity for life history traits in *Drosophila melanogaster*. I. Effect on phenotypic and environmental correlations. Journal of Evolutionary Biology, **6**, 1–16.

Gould, S.J. & Lewontin, R.C. (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London B, **205**, 581–598.

Hegmann, J.P. & Dingle, H. (1982) Phenotypic and genetic covariance structure in milkweed bug life history traits. In: Evolution and genetics of life histories (ed. by H. Dingle and J. P. Hegmann), pp. 177–185. Springer, New York.

Honek, A. (1993) Intraspecific variation in body size and fecundity in insects: a general relationship. Oikos, **66**, 483–492.

Houle, D. (1991) Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution, **45**, 630–648.

Johnson, M.D. (1990) Female size and fecundity in the small carpenter bee, *Ceratina calcarata* (Robertson) (Hymenoptera: Anthophoridae). Journal of the Kansas Entomological Society, **63**, 414–419.

Kasule, F.K. (1991) Associations of fecundity with adult size in the cotton stainer bug *Dysdercus fasciatus*. Heredity, **66**, 281–286.

Klingenberg, C.P. & Spence, J.R. (1993) Heterochrony and allometry: lessons from the water strider genus *Limnoporus*. Evolution, **47**, 1834–1853.

Kusano, T. (1982) Postmetamorphic growth, survival, and age at first reproduction of the salamander, *Hynobius nebulosus tokyoensis* Tago in relation to a consideration on the optimal timing of first reproduction. Researches on Population Ecology, **24**, 329–344.

Lamb, R.J. (1992) Developmental rate of *Acyrthosyphon pisum* (Homoptera: Aphididae) at low temperatures: implications for estimating rate parameters for insects. Environmental Entomology, **21**, 10–19.

Larsson, F.K. (1989) Female longevity and body size as predictors of fecundity and egg length in *Graphosoma lineatum* L. (Hemiptera, Pentatomidae). Deutsche entomologische Zeitschrift, N. F., **36**, 329–334.

Leather, S.R. (1988) Size, reproductive potential and fecundity in insects: things aren't as simple as they seem. Oikos, **51**, 386–389.

Leather, S.R. & Burnand, A.C. (1987) Factors affecting life-history parameters of the pine beauty moth, *Panolis flammea* (D&S): the hidden costs of reproduction. Functional Ecology, **1**, 331–338.

Ludwig, D. & Rowe, L. (1990) Life-history strategies for energy gain and predator avoidance under time constraints. American Naturalist, **135**, 686–707.

Marshall, L.D. (1990) Intraspecific variation in reproductive effort by female *Parapediasia teterrella* (Lepidoptera: Pyralidae) and its relation to body size. Canadian Journal of Zoology, **68**, 44–48.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D. & Wolpert, L. 1985. Developmental constraints and evolution. Quarterly Review of Biology, **60**, 265–287.

McLain, D.K. (1991) Heritability of size: a positive correlate of multiple fitness components in the southern green stink bug (Hemiptera: Pentatomidae). Annals of the Entomological Society of America, **84**, 174–178.

Messina, F.J. (1993) Heritability and 'evolvability' of fitness components in *Callosobruchus maculatus*. Heredity, **71**, 623–629.

Nijhout, H.F. (1979) Stretch-induced molting in *Oncopeltus fasciatus*. Journal of Insect Physiology, **25**, 277–281.

Nijhout, H.F. (1984) Abdominal stretch reception in *Dipetalogaster maximus* (Hemiptera: Reduviidae). Journal of Insect Physiology, **30**, 629–633.

Nijhout, H.F. (1994) Insect Hormones. Princeton University Press, Princeton.

Oreskes, N., Shrader-Frechette, K. & Belitz, K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences. Science, **263**, 641–646.

Orzack, S.H. & Sober, E. (1994) Optimality models and the test of adaptationism. American Naturalist, **143**, 361–380.

Palmer, J.O. & Dingle, H. (1986) Direct and correlated responses to selection among life-history traits in milkweed bugs (*Oncopeltus fasciatus*). Evolution, **40**, 767–777.

Panizzi, A.R. & Saraiva, S.I. (1993) Performance of nymphal and adult southern green stink bug on an overwintering host plant and impact of nymph to adult food-switch. Entomologia experimentalis et applicata, **68**, 109–115.

Partridge, L. & Harvey, P.H. (1988) The ecological context of life history evolution. Science, **241**, 1449–1455.

Peters, R.H. (1983) The Ecological Implications of Body Size. Cambridge University Press, Cambridge, U.K.

Pruess, K.P. (1983) Day-degree methods for pest management. Environmental Entomology, **12**, 613–619.

Reeve, H.K. & Sherman, P.W. (1993) Adaptation and the goals of evolutionary research. Quarterly Review of Biology, **68**, 1–32.

Reiss, M.J. (1989) The Allometry of Growth and Reproduction. Cambridge University Press, Cambridge, U.K.

Roff, D.A. (1980) Optimizing development time in a seasonal environment: the 'ups and downs' of clinal variation. Oecologia (Berlin), **45**, 202–208.

Roff, D.A. (1981) On being the right size. American Naturalist, **118**, 405–422.
Roff, D.A. (1984) The evolution of life history parameters in teleosts. Canadian Journal of Fisheries and Aquatic Sciences, **41**, 989–1000.
Roff, D.A. (1992) The Evolution of Life Histories: Theory and Analysis. Chapman and Hall, New York.
Rowe, L. & Ludwig, D. (1991) Size and timing of metamorphosis in complex life cycles: time constraints and variation. Ecology **72**, 413–427.
Rowe, L. & Scudder, G.G.E. (1990) Reproductive rate and longevity in the waterstrider, *Gerris buenoi*. Canadian Journal of Zoology, **68**, 399–402.
Rubenstein, D.I. (1984) Resource acquisition and alternative mating strategies in water striders. American Zoologist, **24**, 345–353.
SAS Institute, Inc. (1988) SAS/Stat User's Guide. SAS Institute, Inc., Cary, N.C.
Sehnal, F. (1985) Growth and life cycles. In: Comprehensive Insect Physiology, Biochemistry, and Pharmacology (ed. by G. A. Kerkut and L. I. Gilbert), vol. 2, pp. 1–86. Pergamon, Oxford.
Shaw, R.G. (1987) Maximum-likelihood approaches applied to quantitative genetics of natural populations. Evolution, **41**, 812–826.
Slansky, F. (1980) Quantitative food utilization and reproductive allocation by adult milkweed bugs, *Oncopeltus fasciatus*. Physiological Entomology, **5**, 73–86.
Solbreck, C., Olsson, R., Anderson, D.B. & Förare, J. (1989) Size, life history and responses to food shortage in two geographical strains of a seed bug *Lygaeus equestris*. Oikos, **55**, 387–396.
Sota, T. (1993) Performance of *Aedes albopictus* and *A. riversi* larvae (Diptera: Culicidae) in waters that contain tannic acid and decaying leaves: is the treehole species better adapted to treehole water? Annals of the Entomological Society of America, **86**, 450–457.
Spence, J.R. (1981) Experimental analysis of microhabitat selection in water-striders (Heteroptera: Gerridae). Ecology, **62**, 1505–1514.
Spence, J.R. (1983) Pattern and process in co-existence of water-striders (Heteroptera: Gerridae). Journal of Animal Ecology, **52**, 497–511.
Spence, J.R. (1986) Relative impacts of mortality factors in field populations of the waterstrider *Gerris buenoi* Kirkaldy (Heteroptera: Gerridae). Oecologia (Berlin), **70**, 68–76.
Spence, J.R. (1989) The habitat templet and life history strategies of pond skaters (Heteroptera: Gerridae): reproductive potential, phenology, and wing dimorphism. Canadian Journal of Zoology, **67**, 2432–2447.
Spence, J.R., Spence, D.H. & Scudder, G.G.E. (1980) The effect of temperature on growth and development of water strider species (Heteroptera: Gerridae) of central British Columbia and implications for species packing. Canadian Journal of Zoology, **58**, 1813–1820.
Spence, J.R., Zimmermann, M. & J. Wojcicki. (1996) Effects of food limitation and sexual cannibalism on reproductive output of the nusery web spider Dolomedes triton (Araneae: Pisauridae). Oikos, in press.
Stearns, S.C. (1992) The Evolution of Life Histories. Oxford University Press, Oxford.
Stearns, S.C. & Crandall, R.E. (1981) Quantitative predictions of delayed maturity. Evolution, **35**, 455–463.
Stearns, S.C. & Koella, J.C. (1986) The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. Evolution, **40**, 893–913.
van Noordwijk, A.J. & de Jong, G. (1986) Acquisition and allocation of resources: their influence on variation in life history tactics. American Naturalist, **128**, 137–142.
Vepsäläinen, K. (1978) Wing dimorphism and diapause in *Gerris*: determination and adaptive significance. Evolution of Insect Migration and Diapause (ed. by H. Dingle). Springer, New York. Pp. 218–253.
Woodring, J.P. (1983) Control of moulting in the house cricket, *Acheta domesticus*. Journal of Insect Physiology, **29**, 461–464.

Yafuso, M. (1994) Life history traits related to resource partitioning between synhospitalic species of *Colocasiomyia* (Diptera, Drosophilidae) breeding in inflorescences of *Alocasia odora* (Araceae). Ecological Entomology, **19**, 65–73.

8. Synthesis and Outlook

In this doctoral thesis, I have discussed the evolution of ontogenies and the methods for analyzing it. I have illustrated important aspects of these topics in the empirical studies of water striders, examining variation at both macro- and microevolutionary scales.

In my general discussion of allometry and heterochrony (chapters 2 and 3), I have emphasized the differences between the various approaches. Allometry and heterochrony differ fundamentally from each other regarding the dimension of time. Allometry is a purely morphological concept and does not refer to time or age; it deals with the covariation among morphometric characters or with the association of size and geometric shape. In contrast, time is essential for the analysis of heterochrony. Heterochrony and allometry are therefore complimentary parts of a full analysis of evolutionary change in ontogenies. There are additional differences between methodological approaches within both heterochrony and allometry. These are related to the different methods for describing organismal form and its ontogenetic change. Each of these emphasizes a different aspect of ontogeny, and they may lead to different interpretations of the same biological problem; it is therefore essential that authors state clearly which criteria they use.

Heterochrony is notorious for an abundance of jargon (see McKinney and McNamara 1991). Not only are there numerous terms with identical meanings, but here are also terms that can have several very different meanings depending on the context in which they are used. As I have shown for the long-standing controversy about human heterochrony, conflicting interpretations can be resolved by considering the criteria used by different researchers to identify the heterochronic processes involved in evolutionary change. The literature about human evolution demonstrates especially well how key terms such as neoteny and especially retardation can be interpreted in several ways—neoteny because it has been defined formally in different ways in the various existing frameworks, and retardation because it never was defined unequivocally as *either* a slowing of rates *or* a delay of events.

The confusion resulting from this ambiguous terminology calls for a solution. One possibility is to discard all the classical terms with their historical baggage, and use purely descriptive terms instead, as Raff and Wray (1989) did in their proposal for a framework for heterochrony in developmental biology. I am sympathetic to this approach, but I do not think that it will find general acceptance among evolutionary biologists. A less drastic alternative is to call for authors to specify clearly which framework they use to classify heterochronic changes in each particular study, and for readers to be aware of the possible differences when comparing several studies. This leaves the well-established terminology intact, but comparisons among studies will be easier as long as studies explicitly report the data and the criteria used to interpret them in terms of heterochrony.

The comparative study of ontogenies in the six species of the water strider genus *Limnoporus* offers an example of this approach (chapter 4). The study examines both multivariate allometry and heterochrony, and in both respects, the species differ in various ways. Differences among species, or even between sexes within species, are generally complex combinations of the simple "unit

changes." The allometric analysis shows changes of the directions of trajectories, ontogenetic scaling and lateral transposition of trajectories. Some of these changes were subtle; for instance, the changes in directions of trajectories are small enough so that it is still meaningful to compute a common direction for analyses of ontogenetic scaling and lateral transposition. The analysis of heterochrony, using the modified version of the framework of Alberch et al. (1979) for a measure of overall size, shows that differences between species and sexes are the result of combinations of the heterochronic processes. Although the basic patterns of allometry and timing have remained constant over a time scale of at least 50 million generations (Andersen et al. 1993), this study demonstrates some degree of evolutionary flexibility in all aspects of ontogeny.

Phenotypic variation in *Limnoporus canaliculatus* shows the same basic pattern of flexibility that follows a highly structured pattern (chapters 5 and 6). Individual variability is similar in each instar, and it is therefore possible to use a common principal component (CPC) model that assumes a shared pattern of variation in all instars as well as in the between-instar covariances. The use of this model considerably simplifies the data analysis by reducing dimensionality and because the CPCs have simpler patterns of ontogenetic variation. The first two CPCs account for almost all the variation. Instead of having to consider all 24 dimensions resulting from the four morphometric variables in the six instars, which are correlated among each other in a complex way, it is therefore sufficient to carry out separate analyses of variation among stages for the first two CPCs.

Analyses of cumulative values of the first CPC, a "size" component, show that phenotypic variation is mostly concentrated in a single dimension (Fig. 6-2). In contrast, a corresponding analysis of increments suggests that variation in growth increments is not structured in such a way (Fig. 6-4). These patterns agree with the results from the interspecific analysis, where the allometric growth trajectories—cumulative values of the morphometric variables—share many features of variation (see Figs. 4-3 and 4-4), whereas the corresponding growth increments vary markedly among species (Fig. 4-5). To examine the connection of the intraspecific patterns of phenotypic variation to constraints on the evolution of these characters more rigorously, it would be necessary to expand the interspecific comparison so that the evolutionary patterns of covariation among stages could be estimated with confidence. Moreover, a quantitative genetic analysis in at least one, but preferably several species would be desirable to assess the potential of these traits to respond to natural selection.

Instar durations are remarkably variable (Fig. 6-6), indicating that the relationship between intrinsic and extrinsic time varies drastically even within a single population. Correspondingly, instar durations also differ strongly among the sexes and species of the genus *Limnoporus* (Table 4-3). These results suggest that caution is necessary when comparisons are based on age alone, as it is necessary in analyses of organisms with continuous growth, where there is no direct information about ontogenetic stage, as it is provided by the molts for water striders.

The lack of a direct proportionality between size increments and age (Figs. 4-6, 6-9) serves as a warning against the use of size as a proxy for age data. The rate of growth varies both within populations and among species, showing that there can be combined heterochronic changes in both overall size and par-

ticular traits. This is exactly the condition that invalidates the framework of "allometric heterochrony," as I have shown in chapters 3 and 4 with a simple graphical model (Fig. 3-6). This reasoning underscores the necessity to separate heterochrony and allometry clearly.

The field experiment with *Gerris buenoi* also has produced unexpected results regarding the relation between size and age. In this large and multiply replicated experiment, none of the numerous separate estimates (Fig. 7-1, Table 7-1) of the correlation between development time and adult size has the high positive value expected under the current paradigm in life history theory (Roff 1992; Stearns 1992). Moreover, the absence of any correlation between female size and reproductive performance (Fig. 7-2 to 7-4, Table 7-3) suggests that within-population variation in size is unrelated to reproductive fitness and may be selectively neutral altogether. As interspecific differences in morphology and life history of water striders presumably are adaptive (Spence 1981, 1989; Klingenberg and Zimmermann 1992), it will be a challenging task for future studies to delimit the boundaries between domains of mainly neutral variation within populations and of natural selection beyond this range of variation.

It is not clear to which extent stretch-induced molting and other physiological mechanisms known from heteropterans (summary in Nijhout 1994) can account for the relationship between size and development time (see Fig. 7-6). The results from *Limnoporus canaliculatus* (chapter 6) are consistent with this model in later instars, which were also used in the experiments that suggested the physiological mechanism of stretch-induced molting, but the patterns from early instars contradict this mechanism. More extensive experiments will be necessary to resolve this issue, both to measure variation in growth and to understand the physiological mechanisms of growth in early instars.

In this project, I have attempted to integrate developmental and ecological aspects with findings from morphometrics to provide insight into the evolution of ontogenies in water striders. This thesis presents new results about heterochronic processes and life history evolution and shows that water striders are a promising group to investigate these questions further. I have applied new statistical methods specifically designed for the biological problem at hand; possible extensions of this approach make this a promising tool for evolutionary biologists. Several of the results in my study shed new light on previous findings from other organisms. Therefore, I hope that this study will make a useful contribution to the emerging discipline of evolutionary developmental biology.

References

Alberch, P., S. J. Gould, G. F. Oster, D. B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5:296–317.

Andersen, N. M., J. R. Spence, and M. V. H. Wilson. 1993. 50 Million years of structural stasis in water striders (Hemiptera: Gerridae). American Entomologist 39:174–176.

Klingenberg, C. P., and M. Zimmermann. 1992. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of water striders. American Naturalist 140:601–620.

McKinney, M. L., and K. J. McNamara. 1991. Heterochrony: the evolution of ontogeny. Plenum Press, New York.

Nijhout, H. F. 1994. Insect hormones. Princeton University Press, Princeton.

Raff, R. A., and G. A. Wray. 1989. Heterochrony: developmental mechanisms and evolutionary results. Journal of Evolutionary Biology 2:409–434.

Roff, D. A. 1992. The evolution of life histories: theory and analysis. Chapman & Hall, New York.

Spence, J. R. 1981. Experimental analysis of microhabitat selection in water-striders (Heteroptera: Gerridae). Ecology 62:1505–1514.

Spence, J. R. 1989. The habitat templet and life history strategies of pond skaters (Heteroptera: Gerridae): reproductive potential, phenology, and wing dimorphism. Canadian journal of Zoology 67:2432–2447.

Stearns, S. C. 1992. The evolution of life histories. Oxford University Press, Oxford.